STRUCTURAL CHARACTERIZATION AND CONFORMATIONAL DYNAMICS OF CLAMSHELL-LIKE LIGAND-BINDING DOMAINS

by

Tyler Wied

A dissertation submitted to The Johns Hopkins University in conformity with the
requirements for the degree of Doctor of Philosophy.

Baltimore, Maryland

September, 2018

Abstract

Clamshell-like ligand binding domains (LBDs) are a diverse class of receptors present in all clades of life. Ligands bind to a cleft which facilitates a transition from primarily open conformations to closed conformations. The LBDs bind a diverse array of ligands, and exhibit large conformational heterogeneity. This dissertation examines the structure and conformational dynamics of three clamshell-like LBDs, each from a different protein system: ionotropic glutamate receptors (iGluRs), metabotropic glutamate receptors (mGluRs), and ionotropic receptors (IRs). In **Chapter 1**, I review various clamshell-like LBD types and discuss their structure, function, and dynamics. **Chapter 2** introduces molecular dynamics simulation and contains a general protocol for users to set up a similar system on their own. **Chapter 3** reveals that the iGluR GluK2, a kainate receptor, has remarkable conformational flexibility. We also present agreement between theoretical and experimental results. **Chapter 4** examines the conformational free energies of mGluR3 in monomer and dimer states, and reveals an interesting effect of dimerization on the stability of closed state LBDs. Finally, **Chapter 5** discusses initial attempts to express and purify insect odorant receptor IR LBDs. Together, this work advances our understanding of LBDs across a spectrum of receptor families.

Thesis Advisor: Prof. Albert Y Lau

Second Reader: Prof. L. Mario Amzel

Thesis Commitee: Prof. L. Mario Amzel, Prof. Cynthia Wolberger, Prof. Taekjip Ha, Prof. Daniel Leahy, Prof. Jürgen Bosch

Acknowledgments

To Elyse De Laittre, I don't think words can ever properly convey how grateful I am to you. I greatly admire your enormous talent and dedication to whatever you set your mind to. Without your patience, comfort, and support, this would not have been possible.

To my family, especially my parents, Pam and Eric Wied. You raised me to be independent and always support my choices in life. My brother, Bryce, you are determined and adventurous. I envy your travel savviness, perhaps you can teach me one day.

To my advisor, Albert Lau, for supervising my thesis work. Your scientific guidance has been immensely valuable. More importantly, the respect and dignity you give to those around you has not gone unnoticed. Thank you.

To my labmates. John Belcher, your quick wit and bottomless trove of knowledge always kept me on my feet. Alvin Yu, you relentless rigor was indispensable to my own development as a scientist. I also appreciate that you facilitated all of out social activity outside the lab. Afif Bandak, I really enjoyed playing racquetball with you, and I hope I learned at least 0.01% all that you know about finance. Sara Bastos, our talented former laboratory technician. You trained me in protein expression and purification, for which I am very grateful. Remy Yovanno, you are boldly carrying the torch of a new generation of the Lau Lab, and will no doubt teach me how to simulate membrane proteins when the time comes. Alfred Chin, you helped enormously with the projects discussed here; never stop being a "big baller".

To my BCMB classmates and friends, especially Philip Cox, Leah Greenspan, Kapil Ramachandran, Dan Silverman, and Rebecca Keener. To Carolyn Machamer, BCMB director during my time at Hopkins. You were an excellent program director, and you will be missed.

To my thesis committee: Prof. Mario Amzel, Prof. Cynthia Wolberger, Prof. TJ Ha, Prof. Dan Leahy and Prof. Jürgen Bosch for their supportive comments and suggestions.

To past mentors at the University of Wisconsin-Madison: Prof. John W. Moore, Prof. Stephen Gammie, and Prof. M. Thomas Record, Jr. Also, to the post-docs and graduate students who trained me: Dr. D. Benjamin Knowles, Dr. Melissa-Ann Scotti, Dr. C. Michael Saul, and Dr. Grace Lee. Thank you all for training undergraduate students, and introducing me to scientific research.

To past teachers, especially Gordon Kendall, Michael Wittig, and David Haller: you instilled in me a love of learning and the confidence to pursue an academic path.

List of Abbreviations

AMBER: Assisted Model Building with Energy Refinement

AMPA: α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid

ATD: amino terminal domain

CGenFF: CHARMM General Force Field

CHARMM: Chemistry at Harvard Macromolecular Mechanics

COM: center of mass

CRD: cysteine rich domain

CTD: C-terminal domain

DNQX: 6,7-dinitroquinoxaline-2,3-dione

ECD: extracellular domain

FDA: Food and Drug Administration

GAAMP: General Automated Atomic Model Parameterization

GAFF: general AMBER force field

GdnHCl: guanidine hydrochloride

GPCR: G-protein coupled receptor

GROMACS: GROningen MAchine for Chemical Simulations (also, Gromacs Runs One Microsecond At Cannonball Speeds, and 400 more)

GROMOS: GROningen Molecular Simulation

iGluR: ionotropic glutamate receptor

IR: ionotropic receptor

LBD: ligand-binding domain

MD: molecular dynamics

mGluR: metabotropic glutamate receptor

NAMD: Nanoscale Molecular Dynamics

NMA: normal mode analysis

NMDA: N-methyl-D-aspartate

NMR: nuclear magnetic resonance

OPLS: Optimized Potentials for Liquid Simulations

PBP: periplasmic binding protein

PMF: potential of mean force

PCA: principal component analysis

PDB: protein database

QM: quantum mechanics

SAXS: small-angle X-ray scattering

smFRET: single-molecule fluorescence resonance energy transfer

TMD: transmembrane domain

VMD: Visual Molecular Dynamics

WHAM: weighted histogram analysis method

Contents

3 High conformational variability in the GluK2 kainate receptor ligand-binding domain 59

4 Conformational dynamics of a metabotropic glutamate receptor ligand binding domain 98

List of Figures

List of Tables

Chapter 1 - Introduction to clamshell-like ligand-binding domains

This dissertation examines structural and dynamical properties of clamshell-like ligand-binding domains (LBD). Canonically, clamshell-like LBDs are two-state receptors that transition from open to closed conformations in the presence of agonist. This chapter introduces clamshell LBDs generally by way of periplasmic binding proteins, and is followed by an introduction to the specific systems presented in later chapters: 1) ionotropic glutamate receptors, 2) metabotropic glutamate receptors, and 3) insect ionotropic receptors.

1.1 Periplasmic binding proteins

Nutrient sensing is an essential activity for bacterial growth and survival. Gram negative bacteria accomplish this with periplasmic binding proteins (PBP) which couple substrate binding to cellular uptake via inner membrane ABC transport proteins [1]. The PBPs constitute an ancient receptor superfamily, which bind chemically diverse substrates such as amino acids, carbohydrates, ions, vitamins, and peptides [2,3]. The binding and transport system is the largest functional group in the *E coli* genome, consisting of at least 281 protein-coding genes [4, 5].

The 3D structure of a PBP was first solved in 1976 with the X-ray crystal structure of the L-arabinose binding protein to 5 Å resolution [6]. In the subsequent 42 years, over 100 PBP structures have been published [3], generating a wealth of insight into PBP structure and conformational change. Structurally, PBPs are bilobed receptors, with a substrate binding site at the lobe interface, and a hinge connecting the lobes. In the absence of substrate, the PBPs adopt primarily open conformations. Many PBPs, such as the ribose binding protein and the D-allose binding protein, appear to adopt multiple, distinct open conformations in the apo state [2, 7–9]. In the presence of substrate, interactions between substrate and each of the two lobes triggers a hinge bending motion and stabilizes a closed cleft conformation (**Fig. 1-1**). For several PBPs, apo and substrate-bound proteins adopt both closed and open conformations, and substrate binding merely shifts the equilibrium toward closed states [10], suggesting the equilibrium between open and closed states is not all-or-none in apo and ligand-bound forms. The open-to-close transitions of bi-lobed receptors have drawn comparisons to the hinge motion of clamshells [11] and Venus flytraps [1, 12].

Despite the chemical diversity of PBP substrates and low sequence similarity between subtypes, PBPs only adopt one of two general folds [3, 4], referred to as Class I and Class II (**Fig. 1-2**). There are two key structural features that separate Class I from Class II PBPs. One, Class I PBPs typically have at least twelve conserved beta strands (A-L) that occupy equivalent positions in structures, whereas Class II PBPs have only 10 (A-J) conserved beta strands. Also, the J beta strand in Class I PBPs is located in the top lobe, but has been relocated to the bottom lobe in Class II PBPs [4]. These two classes descend from a common ancestor. Class I PBPs most closely resemble the hypothesized progenitor PBP [4]. The architecture for Class I and Class II PBPs is also observed in bacterial cytoplasmic DNA-binding regulatory proteins (e.g. LacI). In eukaryotes, Class I and II PBPs underwent separate gene fusion events with membrane proteins to form functional membrane receptors, where the PBPs now serve as ligand-binding domains (LBDs), e.g. ionotropic glutamate receptors (iGluRs) [13, 14] and Class C GPCRs (e.g. metabotropic glutamate and GABA-B receptors) [12, 15]. Specifically, the folds of Class C GPCR LBDs and iGluR amino terminal domains descend from Class I PBPs, whereas iGluR LBDs descend from Class I PBPs. The functional diversity of clamshell-like binding domains has made them desirable targets for engineering of biosensors [16, 17], novel enzymes [18], and optogenetic tools [19] that have been used to restore a visual response in blind mice [20].

1.2 Ionotropic glutamate receptors

1.2.1 iGluR function

Ionotropic glutamate receptors (iGluRs) are tetrameric ligand-gated ion channels classified primarily into four pharmacologically-derived sub-families: kainate, N-methyl-D-aspartate (NMDA), α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA), and δ [21]. These families are further classified into specific subtypes. In humans, these are GluA1-4 AMPA receptors, GluK1-5 kainate receptors, GluN1, GluN2A-D, GluN3A-B NMDA receptors, and GluRD1,2 receptors. In mammals, iGluRs mediate excitatory neurotransmission at synapses by directly depolarizing post-synaptic neurons in response to neurotransmitter binding [22]. Free glutamate is packaged into vesicles in pre-synaptic neurons, which undergo calcium-dependent exocytosis and release glutamate into the synaptic cleft, resulting in a large increase of synaptic glutamate concentration [23]. Glutamate diffuses across the synaptic cleft and binds iGluRs, which undergo a series of concerted conformational transitions to trigger the opening of the ion channel pore. Na^+, K^+ and sometimes Ca^{2+} cations move through the open pore from the extracellular space into the cell, causing cell depolarization [24]. An action potential is triggered in the post-synaptic neuron if the membrane potential crosses a threshold [25].

Glutamate is the most abundant amino acid in the central nervous system, and is excitatory at over 90% of synapses [26, 27]. Proper iGluR function is required for several higher order cognitive processes such as learning and memory, and its dysfunction has been linked to several neurological disorders, like schizophrenia [28]. Outside the central nervous system, there is iGluR expression in peripheral organs like the pancreas, where AMPA receptors stimulate the release of glucagon [29, 30], and the kidneys, where NMDA receptors contribute to the function of the glomerulus [31]. Broad iGluR expression and functionality throughout the body make iGluRs an attractive pharmacological target for treating disease. They are also an ancient receptor family, and play important physiological roles in many clades of life. For instance, in the plant *Arabidopsis thaliana*, GLUTAMATE RECEPTOR-LIKE genes were found to mediate wound signaling between leaves [32]. To date, iGluRs have been characterized in insects [33], ctenophores [34], bdelloid rotifers [13], and prokaryotic cyanobacteria [35].

1.2.2 iGluR Domains

Structurally, eukaryotic iGluRs are composed of four modular domains: the amino-terminal domain (ATD), the ligand-binding domain (LBD), the transmembrane domain (TMD), and the C-terminal domain (CTD) [36]. The ATD is a clamshell that closely resembles Class I PBPs and Class C GPCR LBDs. The function(s) of the ATDs is relatively unknown and remains the focus of intense research. In NMDA receptors, the ATDs are known to affect the open probability of the ion channel, and act as sites of allosteric regulation by binding molecules in the clamshell cleft [37]. It has been suggested that NMDA ATDs are more susceptible to allosteric modulation than the ATDs of other subtypes because of looser packing between the ATD lower lobes, allowing conformational flexibility and facilitating ATD clamshells to open and close. Additionally, NMDA ATDs are positioned closer to the LBD layer than in AMPA receptors, possibly allowing stronger allosteric communication to be communicated between the ATD and LBD layers in NMDA receptors [38–40]. In non-NMDA receptors, the ATD appears to play a less significant role in altering open channel probabilities, and have a reduced capacity to bind allosteric regulators. This is perhaps due to decreased flexibility of the clamshell cleft, restricted inter-ATD subunit movement, and/or the relatively long linker between the ATD and LBD layer [41]. Recent work has challenged this framework, however. AMPAR ATDs have been shown to possess more conformational flexibility than previously thought, and recent structures depict closer ATD-LBD contact than was present in older structures suggesting the possibility of allostery in AMPAR ATDs [42, 43]. Isolated ATD domains are structurally well-characterized by X-ray crystallography. There are several crystal structures of isolated iGluR ATD for various subtypes, both as homomers (GluA1 [44], GluA2 [45–47], GluA3 [48], GluA4 [43], GluK2 [49], GluK3 [50], GluK5 [50], GluN1 [51, 52], GluN2B [53], GluD1 and GluD2 [54]), and heteromers (GluK2/GluK5 [55], GluN1/GluN2A [56], GluN1/GluN2B [57]). The interface of ATD dimers is reminiscent of mGluR LBD dimers (see **Fig. 4-5**).

The LBDs bind substrate and directly control ion channel pore opening [58]. The LBDs resemble Class II PBPs and are the focus of hundreds of functional and structural studies. Agonist-bound LBD conformations associated with ion channel activation adopt fully-closed conformations, partial agonists primarily induce partially-closed LBD conformations, and apo and antagonist bound LBDs are typically open [59]. The relationship between LBD closure and agonism is routinely observed in the over 300 crystal structures corresponding to full agonism, partial agonism, antagonism, and apo states. There are exceptions to this rule (e.g. the closed cleft of apo GluN3A in [60]), suggesting iGluRs possess conformational flexibility. This dynamic flexibility has been corroborated by a wealth of experimental and computational approaches, including single molecule FRET [61], NMR [62], and molecular dynamics simulation [63]. These approaches have also generated deeper insights into LBD conformational dynamics, channel gating, the thermodynamics of ligand binding and clamshell closure [59], and ligand-binding pathways and kinetics [64, 65].

The majority of iGluR subtypes has an associated crystal structure of its LBD, and occasionally as a homo- or heterodimer, representing a large variety of conformational states: GluA2 ([66–69]), GluA3 ([70,70]), GluA4 ([71]), GluK1 ([67,72–74]), GluK2 ([75]), GluK3 ([75,76]), GluK4 ([77]), GluN1 ([60]), GluN2A ([78]), GluN3A ([60]), and GluD1 ([79]).

The iGluR TMDs share a common ancestor with K^+ channels; all iGluRs likely descend from a gene fusion event between a class II PBP and a K^+ channel [13]. Some non-animal iGluRs retain K^+ selectivity, e.g. GluR0 and AvGluR [13]. Each subunit has four helices (M1-M4), with a re-entrant loop between helices 2 and 3 that contains the selectivity filter, but is topologically inverted compared to K^+ channels [80]. In GluA2, the glutamine to arginine (Q/R) mRNA modification in the re-entrant loop reduces selectivity for calcium [58]. The M3 helix of each subunit forms the ion channel pore, and is directly tethered to the LBD lower lobe [81]. The M4 helix is absent in prokaryotic iGluRs and is important for subunit assembly and trafficking [82].

The CTD is cytoplasmic and responsible for regulating receptor trafficking, long-term potentiation, and channel open probability [58]. These functions are frequently mediated by differential phosphorylation of the CTD [83]. There is considerable variability in CTD length between iGluR subtypes and splicing variants, and the precise roles of the CTD still remain to be resolved. Structurally, CTDs are thought to be mostly intrinsically disordered, and to date there are no 3D structures of any iGluR CTD. Nonetheless, active efforts in this area seek to solve structures of CTDs in complex with other proteins, where CTDs (or CTD sub-regions) are likely to take on preferred configurations.

1.2.3 Tetrameric assembly

NMDA receptors are obligate heterotetramers composed of glutamate-binding GluN2 subunits and glycine-binding GluN1 or GluN3 subunits [84]. AMPA and kainate receptors bind glutamate and form functional homo- and heterotetramers [58]. Three major receptor functional states are observed in electrophysiology: the resting state, active state, and desensitized state (corresponding to a closed channel in the presence of agonist). Each functional state is associated with a set of conformations of the tetramer [58]. Solving the structure of native iGluRs in the active, inactive, and desensitized states is significantly more challenging than isolated domains. Challenges here include the relative difficulty of crystallizing membrane proteins and the transient conformational states associated primarily with activated states, but also other functional states.

Nevertheless, several native iGluR structures have been publish in the past few years. The first structure of a tetrameric iGluR assembly was that of the AMPA receptor GluA2 homotetramer with a closed ion channel pore, and an antagonist bound to the LBD cleft, likely representing an inactive conformation [81]. The tetrameric assembly of this structure adopts a characteristic "Y" shape (**Fig. 1-3** *left*). Subsequent AMPA structures depict pre-open, open, and desensitized conformations [42, 85, 86]. Notably, whereas homomeric inactive GluA2 has a characteristic "Y" shape with significant separation between LBD and ATD layers, the heteromeric inactive GluA2/GluA3 has a much more vertically condensed structure. This opens up the possibility of enhanced allosteric signaling between the ATD and LBD of AMPA heteromers [58]. Structures associated with each of the functional states of the GluA2 homotetramer are shown in **Fig. 1-5**A. Ion channel pore opening appears to be triggered by 1) closure of LBDs, and 2) tilting of the A and C LBD subunits away from the B and D subunits (**Fig. 1-5**B). Structures of the GluA2 desensitized state demonstrate conformational heterogeneity and often substantial disruption of dimer-of-dimer contact in the ATD and LBD layers [86]. This substantial separation is not always present, however [87, 88]. Recent analysis suggests ATD splaying in desensitization could be an artifact of construct design and/or sample preparation, and is unlikely to occur *in vivo* because of crowding [89].

Structures of native NMDA and kainate receptors have also been solved in various conformational states [39, 86, 90–92]. **Fig. 1-3** depicts representative structures of native iGluRs from the AMPA (GluA2), NMDA (GluN1/GluN2A), and Kainate (GluK2) families. The GluA2 and GluN1/GluN2B depicted are inactive assemblies. Both tetramers are globally two-fold symmetric but each layer of the tetramer has distinct symmetry and subunit pairing (see **Fig. 1-4**). The TMD layer is four-fold symmetric in GluA2 and pseudo-four-fold symmetric in GluN1/GluN2B. The LBD layer is a two-fold symmetric dimer-of-dimers where the A/D subunits form one dimer and B/C form the other. The A and C subunits and B and D subunits are equivalent within the LBD layer, respectively. In GluN2B, the 1-β/2B-α forms one dimer, and 2B-β/1-α form the other; 1-α and 1-β subunits and 2B-α and 2B-β are equivalent. The ATD layer has a similar arrangement to the LBD layer, except for a domain dimer partner swap: B/A and C/D are dimer partners in the GluA2 ATD layer, and 2B-β/1-β and 1-α/2B-α are partners in the GluN1/GluN2A ATD layer. In contrast to GluA2 and GluN1/GluN2B, the GluK2 structure shown likely represents a desensitized state. The principal difference is a two-fold to four-fold symmetry transition in the LBD layer. The LBD dimer interfaces are completely disrupted, and the top lobes of the LBDs are oriented outward from the center of the receptor.

1.3 Class C GPCR Ligand binding domains

G-protein coupled receptors (GPCRs) are seven-pass transmembrane receptors that couple extracellular ligand binding to intracellular signaling. Ligand binding triggers a conformational transition that activates guanine exchange factor (GEF) activity which promotes GTP exchange and dissociation of a trimeric G protein from the GPCR. The G protein subunits (α and $\beta\gamma$) in turn activate downstream effector proteins (e.g. Adenylyl cyclase) (**Fig. 1-6**). The GPCRs are a large receptor superfamily, consisting of over 800 genes in humans, or ∼4% of the human protein-coding genome. These genes are further clustered into five classes A-F, which are structurally similar but share <20% sequence identity [93]. Functionally, GPCRs serve in a wide variety of systems: vision, olfaction, gustation, neuronal, and hormone signaling [93]. As such, GPCRs are also the most common drug target: ∼34% of all FDA approved drugs target GPCRs (475 FDA-approved drugs target GPCRs as of July 2017) [94].

The transmembrane domains of GPCRs are relatively well-characterized structurally. To date, there are over 50 atomic structures of unique GPCR proteins, and over 200 atomic structures in total [95]. These studies have yielded invaluable insight into GPCR structure and conformational change. However, most of these structures have removed domains and flexible loops outside the membrane. For example, Class C GPCRs are distinguished by large N-terminal clamshell LBDs that are absent in structures with the TMD, and vice versa. The metabotropic glutamate receptors belong to the Class C GPCRs, and while there are separate structures of mGluR TMDs ([96–98]) and the LBDs ([99–104]), there are no structures with both domains present.

Class C GPCRs are obligate dimers, with dimer interfaces between both the LBD and TMD domains of each subunit (**Fig. 1-7**, *top*). There is also a cysteine rich domain (CRD) between the LBD and TMD. Receptors in this class include calcium-sensing, GABAB , mGluR, RAIG, and taste. Activation of Class C receptors involves ligand binding and LBD closure, which triggers a rotation of the dimer interface (**Fig. 1-7**, *bottom*). The reorientation of the LBD dimer interface is signaled to the TMDs, which trigger downstream events. The inactive state of the LBD dimer is associated with open LBDs, and a dimer interface with separated lower lobes. When agonist binds to the LBD cleft, cleft closure stabilizes the dimer interface rotation to a state where the lower lobes form a dimer interface. In the presence and absence of agonist, LBDs rapidly oscillate between active and inactive conformations. Agonists merely shift the equilibrium such that the receptor has a higher probability of visiting the active conformation. Activation at the LBD layer is presumably transmitted to the TMD layer via the CSD. Re-orientation brings the CSD of each subunit into close proximity, and introducing cysteine locks between CSDs results in a constitutively active receptor. A re-orientation of the dimer interface at the TMD layer is also required for G protein activation.

1.4 Insect ionotropic receptors

A universal property of life is the ability to sense external chemical species. This allows lifeforms to find food, suitable mates, and avoid danger. Whereas mammalian sensory receptors for external chemical stimuli are typically GPCRs, insects primarily utilize ion channels instead [105]. This difference could represent evolutionary happenstance, or reflect the importance of insect reaction time, where signal speed is selected over signal amplification [105]. Chemosensation in insects is divided between olfactory receptors (ORs), pickpocket ion channels (PPKs), transient receptor potential ion channels (TRPs), gustatory receptors (GRs), and ionotropic receptors (IRs) [106]. The ORs and GRs are seven transmembrane ionotropic receptors, and are unrelated to the seven transmembrane GPCRs [107]. Ppks are sodium channels required for water sensing, and TRP channels are six-pass ion channels responsible for thermosensation, the detection of electrophilic compounds, as well as some bitter compounds [108].

The IRs represent a newly characterized class of receptors, with homology to the ionotropic glutamate receptors, but cannot be classified into the NMDA, AMPA, or kainate families [109]. The IRs are a divergent receptor family, composed of over 60 subtypes in *Drosophila*, with 10-70% amino acid sequence identity shared between subunits globally. Compared to the rest of the protein, the LBD region is significantly less conserved on average (<34% shared identity). The putative ligand-binding site lacks conserved residues found in iGluR LBDs, for example, only 31% of IRs possess the α-carboxy-binding L-arginine residue conserved in all iGluRs. The sequence divergence is also reflected in the diversity of ordorants that bind IRs. Despite the large number of odorants recognized by IRs, the types of odorants are limited largely to amines, carboxylic acids, and aldehydes [106, 110–114], reflecting their similarity to iGluRs which bind amino acids. The majority of IRs are expressed in odorant sensory neurons, and are directly responsible for sensing odorants. However, there are exceptions that suggest IRs play roles outside odorant recognition: IR21 and IR25a are required for cold avoidance [115], where IR25a's temperature sensing role also controls the circadian clock [116]; IR64, IR40a, and IR93a play important roles in humidity sensing [117, 118]. IR20a has been proposed to be a taste and pheromone receptor [119]. IRs are of particular interest for the development of novel insect repellents; initially, IR40a was proposed to act as a DEET receptor, however, recent work suggests otherwise [120]. The mechanism of DEET detection in insects remains unresolved.

IRs are thought to be tetrameric ion channels like iGluRs [121]. IR subunits are classified as either odor-specific or co-receptors. Odor-specific IR subunits are selectively expressed in sensory neurons and are responsible for odor recognition. Co-receptor subunits (i.e. IR8a and IR25a) are broadly expressed and required for proper channel assembly, trafficking, and activation [121]. Notably, while odor-specific possess LBDs and TMDs that are related to their iGluR counterparts, their N-terminal regions are unrelated to iGluR ATDs. The co-receptors IR8a and IR25a, however, do have canonical iGluR ATDs (**Fig. 1-8**, *left*). *In vivo* results from a fluorescently-tagged IR84a odor-specific subunit and IR8a co-receptor subunit indicated the stoichiometry in tetrameric assemblies is two co-receptor subunits to two odor-specific receptors, but it is unclear whether this generalizes to all IRs [121]. The proposed arrangement of subunits in the tetrameric assembly shown in **Fig. 1-8** (based on Figure 2B in [114]) was chosen to allow the co-receptor ATDs to occupy the proximal ATD position, where they are able to interact, as opposed to the distal position, where they would not. The model depicted in **Fig. 1-8** is based on the NMDA receptor.

To date, no 3D structural information exists for IRs. Solving the structures of isolated IR ligand-binding domains, amino-terminal domains, and not to mention fully assembled receptors, are necessary to understand the molecular basis of odor recognition and the molecular evolution of iGluRs.

1.5 Overview of remainder of thesis

The remainder of this dissertation presents the conformational and dynamical properties of various clamshell-like LBDs, using both computational and experimental approaches.

Chapter 2 introduces the computational method of molecular dynamics simulation on LBDs presented in Chapters 3 and 4. The chapter begins as an introduction to the principles of molecular dynamics simulation and umbrella sampling, and ends with a practical discussion of how to implement the system.

Chapter 3 presents the energetics of LBD conformational transition in a kainate receptor LBD. Conformational free energy calculations show that apo GluK2 is remarkably flexible. Calculations with disrupted interdomain contacts suggest that residues contribute variably to closed state stability. Comparison of experimental and theoretical SAXS show remarkable agreement, which are a significant improvement over predictions from the crystal structure alone. Finally, we present the first structural characterization of apo GluK2.

Chapter 4 examines the energetics of both monomeric and dimeric mGluR LBDs. Monomeric glutamate-bound mGluR3 exhibits two-state character, but interestingly dimerization to a closed-state LBD partner appears to stabilize closed state conformations. Additionally, the energetics of the LBD dimer interface are examined.

Chapter 5 contains a summary of attempts to experimentally characterize IR LBDs. This covers construct design of various IR subtypes LBDs and expression attempts in both bacterial and baculovirus insect cell system. In brief, isolated IR LBDs appear to be substantially more challenging to express than their iGluR counterparts, although some promising leads were discovered. It is possible iGluR guideposts are misleading, and future studies should potentially re-visit construct design from scratch.

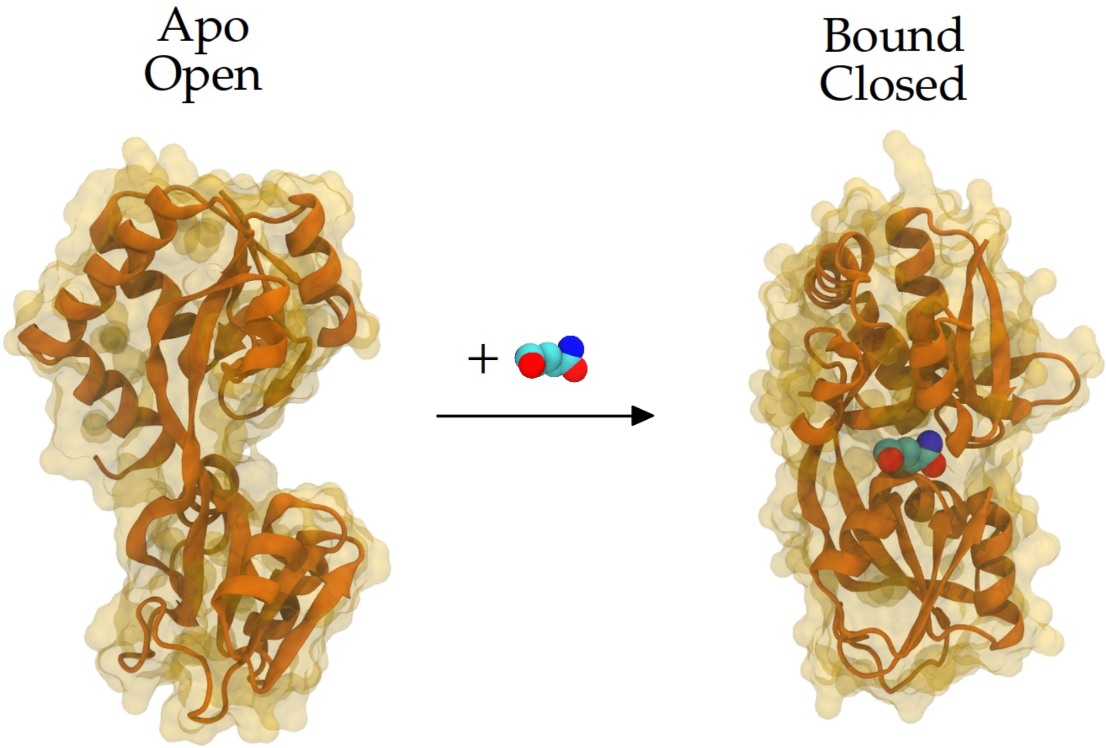


Fig. 1-1: **Ligand binding and clamshell closure in periplasmic binding proteins.** Depiction of the glutamine binding protein transitioning from an apo, open state (*left*, PDB: 1GGG) to a glutamine-bound, closed state (*right*, PDB: 1WDN).

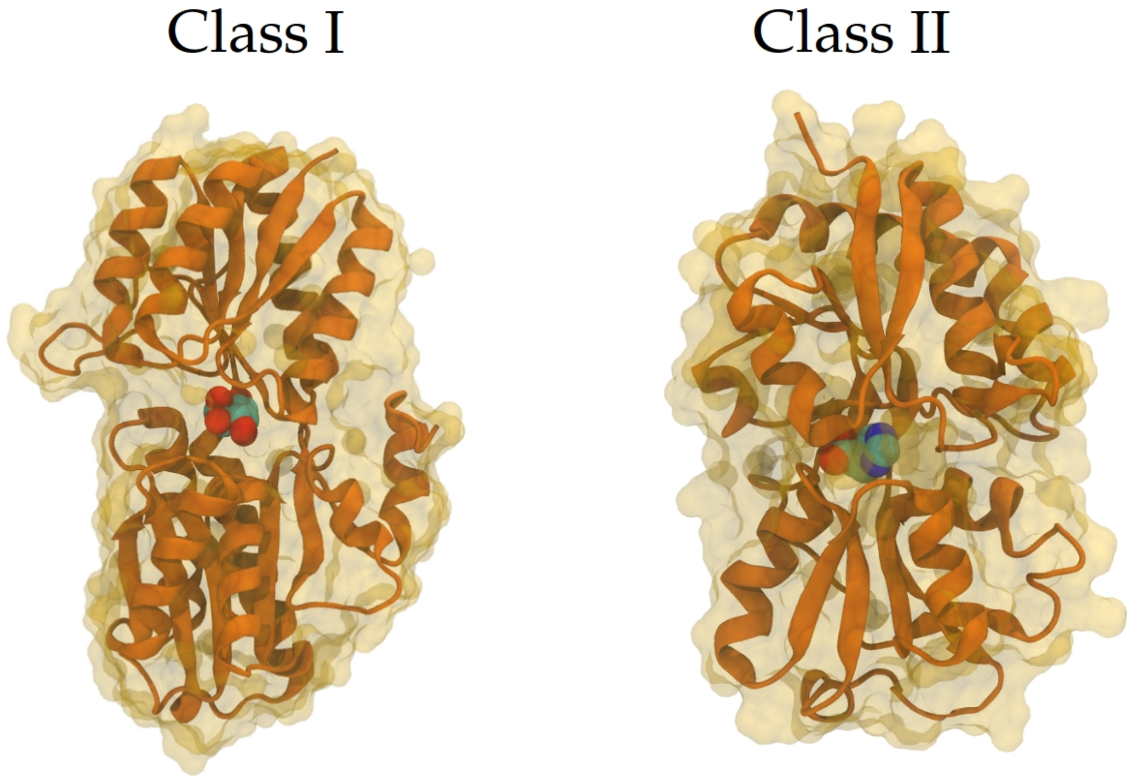


Fig. 1-2: **Structure of class I and II periplasmic binding proteins.** *Left*: L-Arabinose-binding periplasmic protein (PDB: 8ABP), with L-arabinose shown in the binding cleft with VDW representation. *Right*: Histidine-binding periplasmic protein (PDB: 1HSL), with histidine shown in the binding cleft with VDW representation. The architecture of Class I PBPs is also observed in eukaryotic Class C GPCR ligand-binding domains (e.g. mGluRs), as well as iGluR amino terminal domains. The architecture of Class II PBPs resembles the iGluR ligand-binding domains.

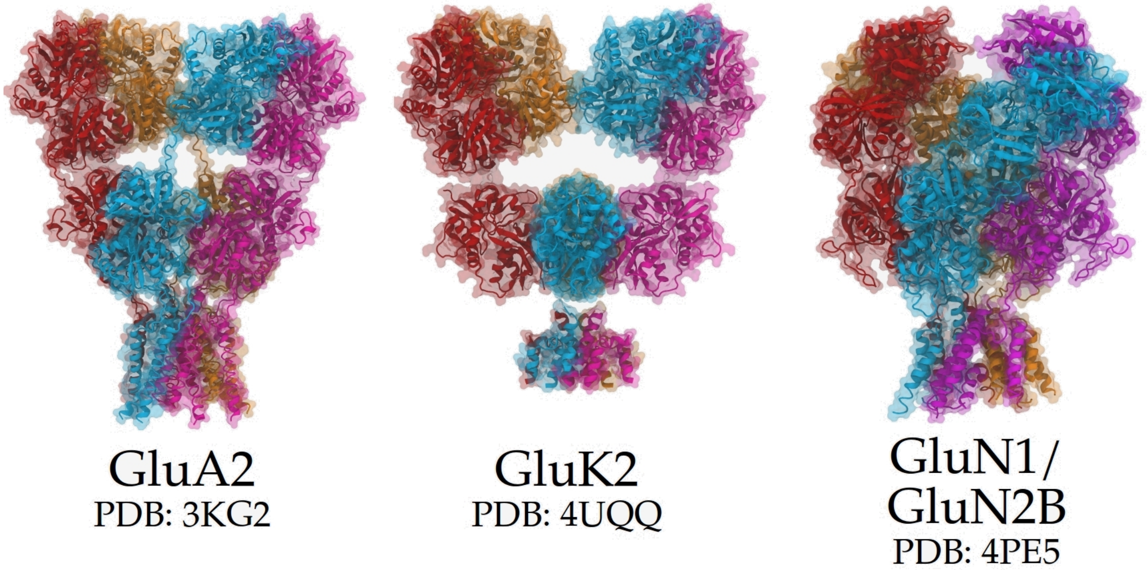


Fig. 1-3: **Ionotropic glutamate receptors.** *Left*: AMPA receptor homotetramer GluA2 bound to antagonist. Center: Kainate receptor homodimer GluK2 bound to agonist, possibly in desensitized conformation. *Right*: NMDA receptor heterotetramer GluN1/GluN2B bound to agonists glycine/glutamate, respectively, in LBD cleft and GluN2B allosteric inhibitor ifenprodil in the ATD.

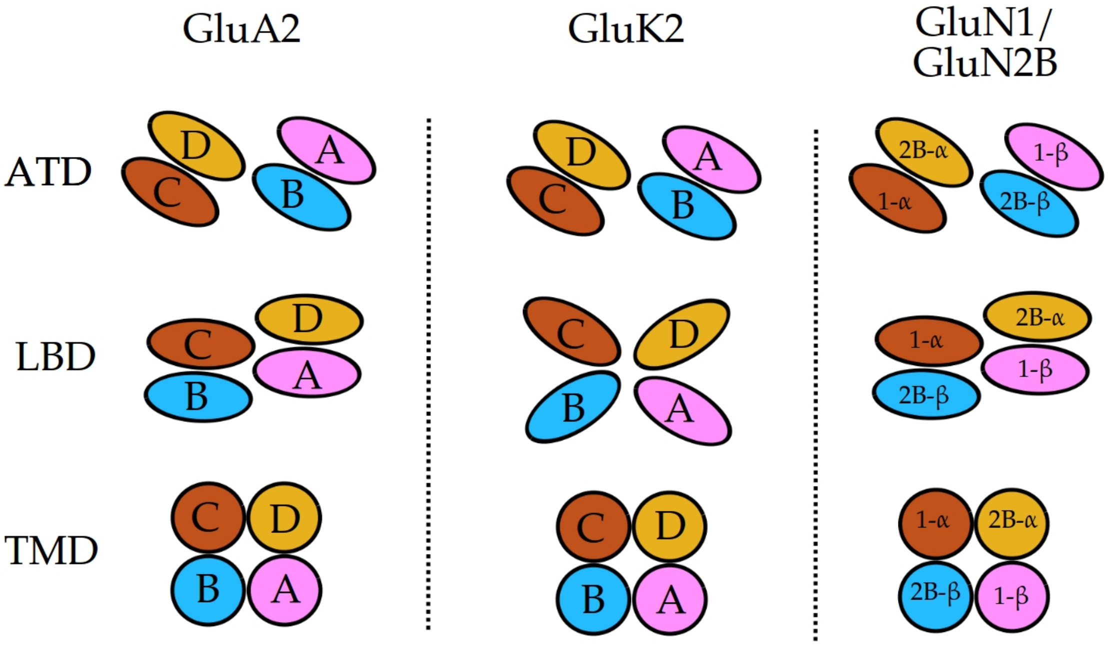


Fig. 1-4: **Symmetry and structural arrangement of iGluR tetramers.** Arrangement of subunits in full-length structures of AMPA GluA2 (*left*), Kainate GluK2 (*center*), and NMDA GluN1/GluN2B (*right*). For GluA2 and GluN1/GluN2B, the crystal structures likely represent inactive states, where the TMDs have either four-fold (GluA2) or pseudo four-fold (GluN1/GluN2B) symmetry, the LBD layer is a two-fold symmetric dimer-of-dimers. Dimer partners are the A/D and B/C subunits in GluA2, and the 1-β/2B-α and 2B-β/1-α subunits in GluN1/GluN2B. The ATD layer is also a two-fold symmetric dimer-of-dimers but with a domain swap from the LBD layer. Dimer partners in the ATD layer are the B/A and C/D subunits in GluA2 and the 2B-β/1-β 1-α/2B-α subunits in GluN1/GluN2B. The arrangement of subunits in GluK2 likely represents a desensitized state, where the TMD and ATD layers maintain the same symmetries seen in NMDA and AMPA receptors, but disrupted dimer interfaces in the LBD layer cause four-fold symmetry instead of two-fold symmetry.

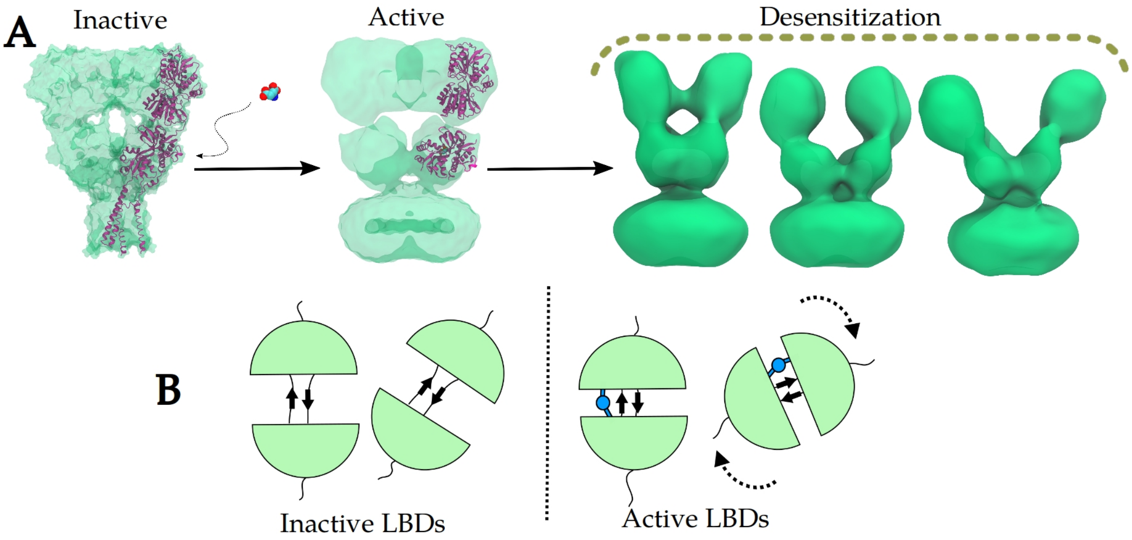


Fig. 1-5: **Functional states of AMPA receptors. A**. Views of the inactive, active, and desensitized states of GluA2. The inactive state has a characteristic "Y" shape, the active state looks broadly similar but has a bulging pore, and the desensitized states exhibit a large variety of conformational heterogeneity characterized by symmetry breaking in the extracellular domains, and large-scale separation of the ECDs. **B**. Cartoon depiction of changes in the LBD layer associated with activation. In the inactive LBD orientation *left*, both LBDs are open and the subunit on the right is at a slight angle to the one on the left. In the transition to the active state *right*, both LBDs close in response to agonist binding, and the subunit on the right undergoes a rotation with respect to the subunit subunit on the left, where the bottom-top axis of the subunit on the right becomes more perpendicular to the bottom-top axis of the subunit on the left.

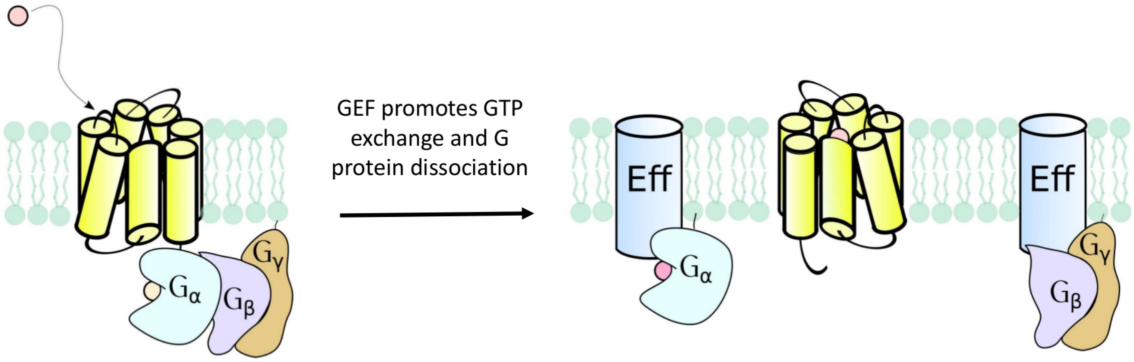


Fig. 1-6: **GPCR activation.** *left* Inactive GPCR (yellow) bound to trimeric G protein ($G_{\alpha\beta\gamma}$). When ligand (red) binds to GPCR, it triggers a conformational change that activates guanine exchange factor (GEF) to exchange out GDP for GTP and promote dissociation of G protein from the GPCR. The G_α and $G_{\beta\gamma}$ subunits separate and are free to activate effector proteins.

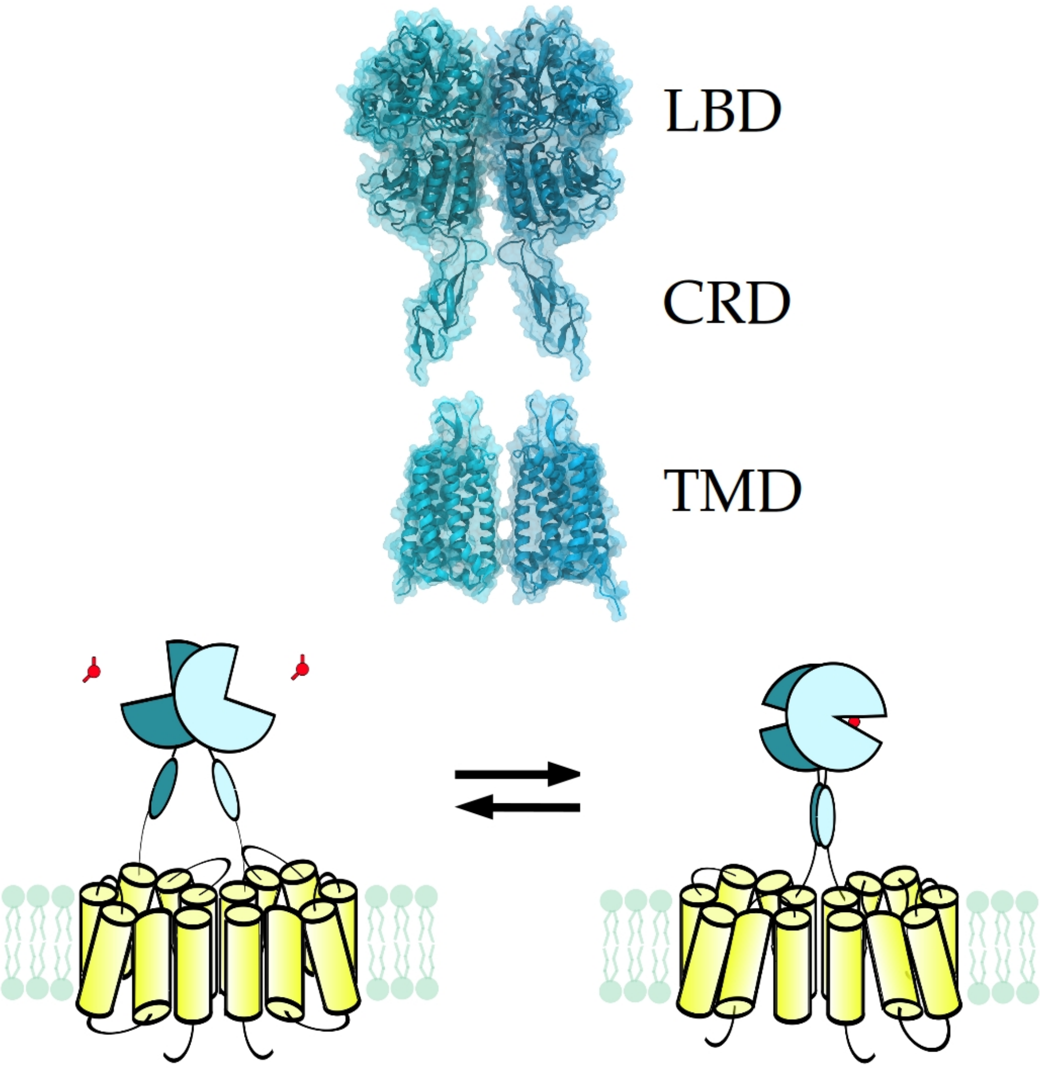
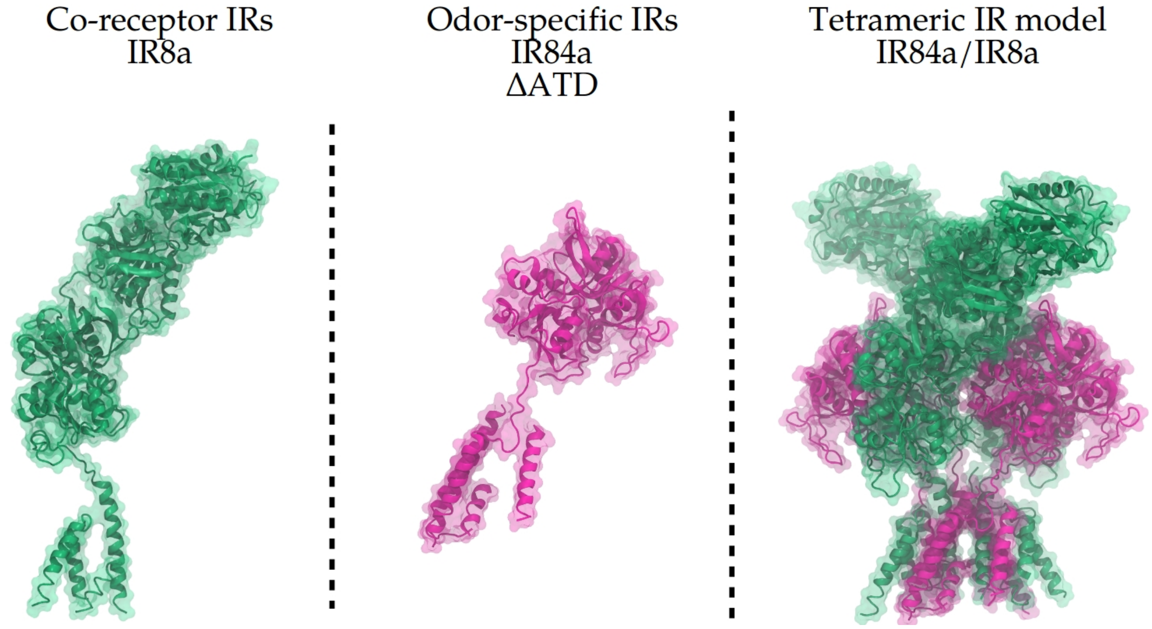


Fig. 1-7: **Structure and activation of metabotropic glutamate receptors.** *Top*, Structural model of a full length mGluR dimer. *Bottom*, Simple model of mGluR activation. Glutamate (red), binds LBDs, which induces cleft closure, and a reorientation of the dimer interface. The CRDs come into close proximity, which is signaled to the TMD which undergoes a conformational transition, which signals to downstream processes.



Co-receptor IRs
IR8a

Odor-specific IRs
IR84a
ΔATD

Tetrameric IR model
IR84a/IR8a

Fig. 1-8: **Structural model of IR tetramer assembly.** Model built using NMDA GluN1/N2B as reference (PDB: 4PE5). *Left*, Co-receptor subunit (e.g. IR8a) with intact ATD. *Center*, odor-specific subunit (e.g. IR84a), lacking an ATD. *Right*, fully assembled tetramer. Co-receptor subunit situated so the ATDs occupy proximal positions (see **Fig. 1-4**, proximal ATD subunits are 2B-α and 2B-β in the GluN1/N2B tetramer).

References

[1] Felder, C. B., Graul, R. C., Lee, A. Y., Merkle, H.-P., and Sadee, W. The Venus flytrap of periplasmic binding proteins: an ancient protein module present in multiple drug receptors. *AAps pharmsci*, 1(2):7–26, 1999.

[2] Borrok, M. J., Zhu, Y., Forest, K. T., and Kiessling, L. L. Structure-based design of a periplasmic binding protein antagonist that prevents domain closure. *ACS chemical biology*, 4(6):447–456, 2009.

[3] Dwyer, M. A. and Hellinga, H. W. Periplasmic binding proteins: a versatile superfamily for protein engineering. *Current opinion in structural biology*, 14(4):495–504, 2004.

[4] Fukami-Kobayashi, K., Tateno, Y., and Nishikawa, K. Domain dislocation: a change of core structure in periplasmic binding proteins in their evolutionary history. *Journal of molecular biology*, 286(1):279–290, 1999.

[5] Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., et al. The complete genome sequence of Escherichia coli K-12. *Science*, 277(5331):1453–1462, 1997.

[6] Phillips, G. N., Mahajan, V. K., Siu, A., and Quiocho, F. A. Structure of L-arabinose-binding protein from Escherichia coli at 5 Å resolution and preliminary results at 3.5 Å. *Proceedings of the National Academy of Sciences*, 73(7):2186–2190, 1976.

[7] Björkman, A. J. and Mowbray, S. L. Multiple open forms of ribose-binding protein trace the path of its conformational change. *Journal of molecular biology*, 279(3):651–664, 1998.

[8] Magnusson, U., Chaudhuri, B. N., Ko, J., Park, C., Jones, T. A., and Mowbray, S. L. Hinge-bending motion of D-allose binding protein from Escherichia coli: three open conformations. *Journal of Biological Chemistry*, 2002.

[9] Magnusson, U., Salopek-Sondi, B., Luck, L. A., and Mowbray, S. L. X-ray structures of the leucine-binding protein illustrate conformational changes and the basis of ligand specificity. *Journal of Biological Chemistry*, 279(10):8747–8752, 2004.

[10] Messina, T. C. and Talaga, D. S. Protein free energy landscapes remodeled by ligand binding. *Biophysical journal*, 93(2):579–585, 2007.

[11] Zhu, S., Stroebel, D., Yao, C. A., Taly, A., and Paoletti, P. Allosteric signaling and dynamics of the clamshell-like NMDA receptor GluN1 N-terminal domain. *Nature Structural and Molecular Biology*, 20(4):477, 2013.

[12] Cao, J., Huang, S., Qian, J., Huang, J., Jin, L., Su, Z., Yang, J., and Liu, J. Evolution of the class C GPCR Venus flytrap modules involved positive selected functional divergence. *BMC evolutionary biology*, 9(1):67, 2009.

[13] Janovjak, H., Sandoz, G., and Isacoff, E. A modern ionotropic glutamate receptor with a K^+ selectivity signature sequence. *Nature communications*, 2:232, 2011.

[14] Wo, Z. G. and Oswald, R. E. Unraveling the modular design of glutamate-gated ion channels. *Trends in neurosciences*, 18(4):161–168, 1995.

[15] O'Hara, P. J., Sheppard, P. O., Thógersen, H., Venezia, D., Haldeman, B. A., McGrane, V., Houamed, K. M., Thomsen, C., Gilbert, T. L., and Mulvihill, E. R. The ligand-binding domain in metabotropic glutamate receptors is related to bacterial periplasmic binding proteins. *Neuron*, 11(1):41–52, 1993.

[16] Hellinga, H. W. and Marvin, J. S. Protein engineering and the development of generic biosensors. *Trends in biotechnology*, 16(4):183–189, 1998.

[17] Ko, W., Kim, S., and Lee, H. S. Engineering a periplasmic binding protein for amino acid sensors with improved binding properties. *Organic & biomolecular chemistry*, 15(41):8761–8769, 2017.

[18] Benson, D. E., Haddy, A. E., and Hellinga, H. W. Converting a maltose receptor into a nascent binuclear copper oxygenase by computational design. *Biochemistry*, 41(9):3262–3269, 2002.

[19] Volgraf, M., Gorostiza, P., Numano, R., Kramer, R. H., Isacoff, E. Y., and Trauner, D. Allosteric control of an ionotropic glutamate receptor with an optical switch. *Nature chemical biology*, 2(1):47, 2006.

[20] Caporale, N., Kolstad, K. D., Lee, T., Tochitsky, I., Dalkara, D., Trauner, D., Kramer, R., Dan, Y., Isacoff, E. Y., and Flannery, J. G. LiGluR restores visual responses in rodent models of inherited blindness. *Molecular Therapy*, 19(7):1212–1219, 2011.

[21] Mayer, M. L. Structural biology of glutamate receptor ion channel complexes. *Current Opinion in Structural Biology*, 41:119–127, December 2016.

[22] Niciu, M. J., Kelmendi, B., and Sanacora, G. Overview of glutamatergic neurotransmission in the nervous system. *Pharmacology Biochemistry and Behavior*, 100(4):656–664, 2012.

[23] Omote, H., Miyaji, T., Juge, N., and Moriyama, Y. Vesicular neurotransmitter transporter: bioenergetics and regulation of glutamate transport. *Biochemistry*, 50(25):5558–5565, 2011.

[24] Dingledine, R., Borges, K., Bowie, D., and Traynelis, S. F. The glutamate receptor ion channels. *Pharmacological reviews*, 51(1):7–62, 1999.

[25] Stuart, G., Spruston, N., Sakmann, B., and Häusser, M. Action potential initiation and backpropagation in neurons of the mammalian CNS. *Trends in neurosciences*, 20(3):125–131, 1997.

[26] Sapolsky, R. M. *Biology and human behavior: The neurological origins of individuality.* Teaching Company, 2005.

[27] Zhou, Y. and Danbolt, N. Glutamate as a neurotransmitter in the healthy brain. *Journal of neural transmission*, 121(8):799–817, 2014.

[28] Rubio, M. D., Drummond, J. B., and Meador-Woodruff, J. H. Glutamate receptor abnormalities in schizophrenia: implications for innovative treatments. *Biomolecules & therapeutics*, 20(1):1, 2012.

[29] Weaver, C. D., Yao, T. L., Powers, A. C., and Verdoorn, T. A. Differential expression of glutamate receptor subtypes in rat pancreatic islets. *Journal of Biological Chemistry*, 271(22):12977–12984, 1996.

[30] Bertrand, G., Gross, R., Puech, R., Loubatières-Mariani, M.-M., and Bockaert, J. Glutamate stimulates glucagon secretion via an excitatory amino acid receptor of the AMPA subtype in rat pancreas. *European journal of pharmacology*, 237(1):45–50, 1993.

[31] Giardino, L., Armelloni, S., Corbelli, A., Mattinzoli, D., Zennaro, C., Guerrot, D., Tourrel, F., Ikehata, M., Li, M., Berra, S., et al. Podocyte glutamatergic signaling contributes to the function of the glomerular filtration barrier. *Journal of the American Society of Nephrology*, 20(9):1929–1940, 2009.

[32] Meinke, D. W., Cherry, J. M., Dean, C., Rounsley, S. D., and Koornneef, M. Arabidopsis thaliana: a model plant for genome analysis. *Science*, 282(5389):662–682, 1998.

[33] Li, Y., Dharkar, P., Han, T.-H., Serpe, M., Lee, C.-H., and Mayer, M. L. Novel functional properties of Drosophila CNS glutamate receptors. *Neuron*, 92(5):1036–1048, 2016.

[34] Alberstein, R., Grey, R., Zimmet, A., Simmons, D. K., and Mayer, M. L. Glycine activated ion channel subunits encoded by ctenophore glutamate receptor genes. *Proceedings of the National Academy of Sciences*, 112(44):E6048–E6057, 2015.

[35] Chen, G.-Q., Cui, C., Mayer, M. L., and Gouaux, E. Functional characterization of a potassium-selective prokaryotic glutamate receptor. *Nature*, 402(6763):817, 1999.

[36] Mayer, M. L. Emerging models of glutamate receptor ion channel structure and function. *Structure*, 19(10):1370–1380, 2011.

[37] Zhu, S. and Paoletti, P. Allosteric modulators of NMDA receptors: multiple sites and mechanisms. *Current opinion in pharmacology*, 20:14–23, 2015.

[38] Mony, L., Kew, J. N., Gunthorpe, M. J., and Paoletti, P. Allosteric modulators of NR2B-containing NMDA receptors: molecular mechanisms and therapeutic potential. *British journal of pharmacology*, 157(8):1301–1317, 2009.

[39] Karakas, E. and Furukawa, H. Crystal structure of a heterotetrameric NMDA receptor ion channel. *Science*, 344(6187):992–997, 2014.

[40] Krieger, J., Bahar, I., and Greger, I. H. Structure, dynamics, and allosteric potential of ionotropic glutamate receptor N-terminal domains. *Biophysical journal*, 109(6):1136–1148, 2015.

[41] Furukawa, H. Structure and function of glutamate receptor amino terminal domains. *The Journal of physiology*, 590(1):63–72, 2012.

[42] Herguedas, B., García-Nafría, J., Cais, O., Fernández-Leiro, R., Krieger, J., Ho, H., and Greger, I. H. Structure and organization of heteromeric AMPA-type glutamate receptors. *Science*, 352(6285):aad3873, 2016.

[43] Dutta, A., Shrivastava, I. H., Sukumaran, M., Greger, I. H., and Bahar, I. Comparative dynamics of nmda-and ampa-glutamate receptor n-terminal domains. *Structure*, 20(11):1838–1849, 2012.

[44] Yao, G., Zong, Y., Gu, S., Zhou, J., Xu, H., Mathews, I. I., and Jin, R. Crystal structure of the glutamate receptor GluA1 N-terminal domain. *Biochemical Journal*, 438(2):255–263, 2011.

[45] Jin, R., Singh, S. K., Gu, S., Furukawa, H., Sobolevsky, A. I., Zhou, J., Jin, Y., and Gouaux, E. Crystal structure and association behaviour of the GluR2 amino-terminal domain. *The EMBO Journal*, 28(12):1812–1823, 2009.

[46] Clayton, A., Siebold, C., Gilbert, R. J., Sutton, G. C., Harlos, K., McIlhinney, R. J., Jones, E. Y., and Aricescu, A. R. Crystal structure of the GluR2 amino-terminal domain provides insights into the architecture and assembly of ionotropic glutamate receptors. *Journal of molecular biology*, 392(5):1125–1132, 2009.

[47] Rossmann, M., Sukumaran, M., Penn, A. C., Veprintsev, D. B., Babu, M. M., and Greger, I. H. Subunit-selective N-terminal domain associations organize the formation of AMPA receptor heteromers. *The EMBO Journal*, 30(5):959–971, 2011.

[48] Sukumaran, M., Rossmann, M., Shrivastava, I., Dutta, A., Bahar, I., and Greger, I. H. Dynamics and allosteric potential of the AMPA receptor N-terminal domain. *The EMBO journal*, 30(5):972–982, 2011.

[49] Kumar, J., Schuck, P., Jin, R., and Mayer, M. L. The N-terminal domain of GluR6-subtype glutamate receptor ion channels. *Nature structural & molecular biology*, 16(6):631, 2009.

[50] Kumar, J. and Mayer, M. L. Crystal structures of the glutamate receptor ion channel GluK3 and GluK5 amino-terminal domains. *Journal of molecular biology*, 404(4):680–696, 2010.

[51] Farina, A. N., Blain, K. Y., Maruo, T., Kwiatkowski, W., Choe, S., and Nakagawa, T. Separation of domain contacts is required for heterotetrameric assembly of functional NMDA receptors. *Journal of Neuroscience*, 31(10):3565–3579, 2011.

[52] Karakas, E., Simorowski, N., and Furukawa, H. Subunit arrangement and phenylethanolamine binding in GluN1/GluN2B NMDA receptors. *Nature*, 475(7355):249, 2011.

[53] Karakas, E., Simorowski, N., and Furukawa, H. Structure of the zinc-bound amino-terminal domain of the nmda receptor nr2b subunit. *The EMBO journal*, 28(24):3910–3920, 2009.

[54] Elegheert, J., Kakegawa, W., Clay, J. E., Shanks, N. F., Behiels, E., Matsuda, K., Kohda, K., Miura, E., Rossmann, M., Mitakidis, N., et al. Structural basis for integration of GluD receptors within synaptic organizer complexes. *Science*, 353(6296):295–299, 2016.

[55] Kumar, J., Schuck, P., and Mayer, M. L. Structure and assembly mechanism for heteromeric kainate receptors. *Neuron*, 71(2):319–331, 2011.

[56] Romero-Hernandez, A., Simorowski, N., Karakas, E., and Furukawa, H. Molecular basis for subtype specificity and high-affinity zinc inhibition in the GluN1-GluN2A NMDA receptor amino-terminal domain. *Neuron*, 92(6):1324–1336, 2016.

[57] Song, X., Jensen, M. Ø., Jogini, V., Stein, R. A., Lee, C.-H., Mchaourab, H. S., Shaw, D. E., and Gouaux, E. Mechanism of NMDA receptor channel block by MK-801 and memantine. *Nature*, 556(7702):515, 2018.

[58] Greger, I. H., Watson, J. F., and Cull-Candy, S. G. Structural and functional architecture of AMPA-type glutamate receptors and their auxiliary proteins. *Neuron*, 94(4):713–730, 2017.

[59] Lau, A. Y. and Roux, B. The hidden energetics of ligand binding and activation in a glutamate receptor. *Nature Structural & Molecular Biology*, 18(3):283–287, March 2011.

[60] Yao, Y., Belcher, J., Berger, A. J., Mayer, M. L., and Lau, A. Y. Conformational analysis of NMDA receptor GluN1, GluN2, and GluN3 ligand-binding domains

reveals subtype-specific characteristics. *Structure*, 21(10):1788–1799, October 2013.

[61] Cooper, D. R., Dolino, D. M., Jaurich, H., Shuang, B., Ramaswamy, S., Nurik, C. E., Chen, J., Jayaraman, V., and Landes, C. F. Conformational transitions in the glycine-bound gluN1 NMDA receptor LBD via single-molecule FRET. *Biophysical journal*, 109(1):66–75, 2015.

[62] Ahmed, A. H., Loh, A. P., Jane, D. E., and Oswald, R. E. Dynamics of the S1S2 glutamate binding domain of GluR2 measured using 19F NMR spectroscopy. *Journal of Biological Chemistry*, 2007.

[63] Lau, A. Y. and Roux, B. The free energy landscapes governing conformational changes in a glutamate receptor ligand-binding domain. *Structure*, 15(10):1203–1214, 2007.

[64] Yu, A., Salazar, H., Plested, A. J., and Lau, A. Y. Neurotransmitter funneling optimizes glutamate receptor kinetics. *Neuron*, 97(1):139–149, 2018.

[65] Yu, A. and Lau, A. Y. Glutamate and glycine binding to the NMDA receptor. *Structure*, 2018.

[66] Krintel, C., Francotte, P., Pickering, D. S., Juknaitė, L., Pøhlsgaard, J., Olsen, L., Frydenvang, K., Goffin, E., Pirotte, B., and Kastrup, J. S. Enthalpy-entropy compensation in the binding of modulators at ionotropic glutamate receptor GluA2. *Biophysical journal*, 110(11):2397–2406, 2016.

[67] Krogsgaard-Larsen, N., Storgaard, M., Møller, C., Demmer, C. S., Hansen, J., Han, L., Monrad, R. N., Nielsen, B., Tapken, D., Pickering, D. S., et al. Structure–activity relationship study of ionotropic glutamate receptor antagonist (2S, 3R)-3-(3-carboxyphenyl) pyrrolidine-2-carboxylic acid. *Journal of medicinal chemistry*, 58(15):6131–6150, 2015.

[68] Szymanska, E., Frydenvang, K., Contreras-Sanz, A., Pickering, D. S., Frola, E., Serafimoska, Z., Nielsen, B., Kastrup, J. S., and Johansen, T. N. A new

phenylalanine derivative acts as an antagonist at the AMPA receptor GluA2 and introduces partial domain closure: synthesis, resolution, pharmacology, and crystal structure. *Journal of medicinal chemistry*, 54(20):7289–7298, 2011.

[69] Ahmed, A. H., Hamada, M., Shinada, T., Ohfune, Y., Weerasinghe, L., Garner, P. P., and Oswald, R. E. The structure of (-)-kaitocephalin bound to the ligand binding domain of the (s)-α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA)/glutamate receptor, GluA2. *Journal of Biological Chemistry*, 287(49):41007–41013, 2012.

[70] Holley, S. M., Ahmed, A. H., Srinivasan, J., Murthy, S. E., Weiland, G. A., Oswald, R. E., and Nowak, L. M. The loss of an electrostatic contact unique to AMPA receptor ligand binding domain 2 slows channel activation. *Biochemistry*, 51(19):4015–4027, 2012.

[71] Birdsey-Benson, A., Gill, A., Henderson, L. P., and Madden, D. R. Enhanced efficacy without further cleft closure: reevaluating twist as a source of agonist efficacy in AMPA receptors. *Journal of Neuroscience*, 30(4):1463–1470, 2010.

[72] Larsen, A. P., Fièvre, S., Frydenvang, K., Francotte, P., Pirotte, B., Karstrup, J. S., and Mulle, C. Identification and structure-function study of positive allosteric modulators of kainate receptors. *Molecular pharmacology*, pages mol–116, 2017.

[73] Krogsgaard-Larsen, N., Delgar, C. G., Koch, K., Brown, P. M., Møller, C., Han, L., Huynh, T. H., Hansen, S. W., Nielsen, B., Bowie, D., et al. Design and synthesis of a series of l-trans-4-substituted prolines as selective antagonists for the ionotropic glutamate receptors including functional and X-ray crystallographic studies of new subtype selective kainic acid receptor subtype 1 (GluK1) antagonist (2S, 4R)-4-(2-Carboxyphenoxy) pyrrolidine-2-carboxylic acid. *Journal of medicinal chemistry*, 60(1):441–457, 2016.

[74] Demmer, C. S., Møller, C., Brown, P. M., Han, L., Pickering, D. S., Nielsen, B., Bowie, D., Frydenvang, K., Kastrup, J. S., and Bunch, L. Binding mode of

an α-amino acid-linked quinoxaline-2, 3-dione analogue at glutamate receptor subtype GluK1. *ACS chemical neuroscience*, 6(6):845–854, 2015.

[75] Mayer, M. L. Crystal structures of the GluR5 and GluR6 ligand binding cores: molecular mechanisms underlying kainate receptor selectivity. *Neuron*, 45(4):539–552, 2005.

[76] Møllerud, S., Pinto, A., Marconi, L., Frydenvang, K., Thorsen, T. S., Laulumaa, S., Venskutonyte, R., Winther, S., Moral, A. M. C., Tamborini, L., et al. Structure and affinity of two bicyclic glutamate analogues at AMPA and kainate receptors. *ACS chemical neuroscience*, 8(9):2056–2064, 2017.

[77] Kristensen, O., Kristensen, L. B., Møllerud, S., Frydenvang, K., Pickering, D. S., and Kastrup, J. S. The structure of a high-affinity kainate receptor: GluK4 ligand-binding domain crystallized with kainate. *Structure*, 24(9):1582–1589, 2016.

[78] Furukawa, H., Singh, S. K., Mancusso, R., and Gouaux, E. Subunit arrangement and function in NMDA receptors. *Nature*, 438(7065):185, 2005.

[79] Naur, P., Hansen, K. B., Kristensen, A. S., Dravid, S. M., Pickering, D. S., Olsen, L., Vestergaard, B., Egebjerg, J., Gajhede, M., Traynelis, S. F., et al. Ionotropic glutamate-like receptor $\delta 2$ binds D-serine and glycine. *Proceedings of the National Academy of Sciences*, 104(35):14116–14121, 2007.

[80] Traynelis, S. F., Wollmuth, L. P., McBain, C. J., Menniti, F. S., Vance, K. M., Ogden, K. K., Hansen, K. B., Yuan, H., Myers, S. J., and Dingledine, R. Glutamate receptor ion channels: structure, regulation, and function. *Pharmacological Reviews*, 62(3):405–496, September 2010.

[81] Sobolevsky, A. I., Rosconi, M. P., and Gouaux, E. X-ray structure, symmetry and mechanism of an AMPA-subtype glutamate receptor. *Nature*, 462(7274):745–756, December 2009.

[82] Salussolia, C. L., Gan, Q., Kazi, R., Singh, P., Allopenna, J., Furukawa, H., and Wollmuth, L. P. A eukaryotic specific transmembrane segment is required for tetramerization in ampa receptors. *Journal of Neuroscience*, 33(23):9840–9845, 2013.

[83] Tingley, W. G., Roche, K. W., Thompson, A. K., and Huganir, R. L. Regulation of NMDA receptor phosphorylation by alternative splicing of the C-terminal domain. *Nature*, 364(6432):70, 1993.

[84] Lee, C.-H. and Gouaux, E. Amino terminal domains of the NMDA receptor are organized as local heterodimers. *PLoS One*, 6(4):e19180, 2011.

[85] Dürr, K. L., Chen, L., Stein, R. A., De Zorzi, R., Folea, I. M., Walz, T., Mchaourab, H. S., and Gouaux, E. Structure and dynamics of AMPA receptor GluA2 in resting, pre-open, and desensitized states. *Cell*, 158(4):778–792, 2014.

[86] Meyerson, J. R., Kumar, J., Chittori, S., Rao, P., Pierson, J., Bartesaghi, A., Mayer, M. L., and Subramaniam, S. Structural mechanism of glutamate receptor activation and desensitization. *Nature*, 514(7522):328–334, October 2014.

[87] Twomey, E. C., Yelshanskaya, M. V., Grassucci, R. A., Frank, J., and Sobolevsky, A. I. Structural bases of desensitization in ampa receptor-auxiliary subunit complexes. *Neuron*, 94(3):569–580, 2017.

[88] Twomey, E. C., Yelshanskaya, M. V., Grassucci, R. A., Frank, J., and Sobolevsky, A. I. Channel opening and gating mechanism in ampa-subtype glutamate receptors. *Nature*, 549(7670):60, 2017.

[89] Twomey, E. C. and Sobolevsky, A. I. Structural mechanisms of gating in ionotropic glutamate receptors. *Biochemistry*, 57(3):267–276, 2017.

[90] Lee, C.-H., Lü, W., Michel, J. C., Goehring, A., Du, J., Song, X., and Gouaux, E. NMDA receptor structures reveal subunit arrangement and pore architecture. *Nature*, 511(7508):191, 2014.

[91] Tajima, N., Karakas, E., Grant, T., Simorowski, N., Diaz-Avalos, R., Grigorieff, N., and Furukawa, H. Activation of NMDA receptors and the mechanism of inhibition by ifenprodil. *Nature*, 534(7605):63, 2016.

[92] Zhu, S., Stein, R. A., Yoshioka, C., Lee, C.-H., Goehring, A., Mchaourab, H. S., and Gouaux, E. Mechanism of NMDA receptor inhibition and activation. *Cell*, 165(3):704–714, 2016.

[93] Katritch, V., Cherezov, V., and Stevens, R. C. Structure-function of the G protein–coupled receptor superfamily. *Annual review of pharmacology and toxicology*, 53:531–556, 2013.

[94] Hauser, A. S., Attwood, M. M., Rask-Andersen, M., Schiöth, H. B., and Gloriam, D. E. Trends in GPCR drug discovery: new agents, targets and indications. *Nature Reviews Drug Discovery*, 16(12):829, 2017.

[95] Thal, D. M., Vuckovic, Z., Draper-Joyce, C. J., Liang, Y.-L., Glukhova, A., Christopoulos, A., and Sexton, P. M. Recent advances in the determination of G protein-coupled receptor structures. *Current opinion in structural biology*, 51:28–34, 2018.

[96] Wu, H., Wang, C., Gregory, K. J., Han, G. W., Cho, H. P., Xia, Y., Niswender, C. M., Katritch, V., Meiler, J., Cherezov, V., et al. Structure of a class C GPCR metabotropic glutamate receptor 1 bound to an allosteric modulator. *Science*, 344(6179):58–64, 2014.

[97] Doré, A. S., Okrasa, K., Patel, J. C., Serrano-Vega, M., Bennett, K., Cooke, R. M., Errey, J. C., Jazayeri, A., Khan, S., Tehan, B., et al. Structure of class C GPCR metabotropic glutamate receptor 5 transmembrane domain. *Nature*, 511(7511):557, 2014.

[98] Christopher, J. A., Aves, S. J., Bennett, K. A., Doré, A. S., Errey, J. C., Jazayeri, A., Marshall, F. H., Okrasa, K., Serrano-Vega, M. J., Tehan, B. G.,

et al. Fragment and structure-based drug discovery for a class C GPCR: discovery of the mGlu5 negative allosteric modulator HTL14242 (3-chloro-5-[6-(5-fluoropyridin-2-yl) pyrimidin-4-yl] benzonitrile). *Journal of medicinal chemistry*, 58(16):6653–6664, 2015.

[99] Kunishima, N., Shimada, Y., Tsuji, Y., Sato, T., Yamamoto, M., Kumasaka, T., Nakanishi, S., Jingami, H., and Morikawa, K. Structural basis of glutamate recognition by a dimeric metabotropic glutamate receptor. *Nature*, 407(6807):971, 2000.

[100] Tsuchiya, D., Kunishima, N., Kamiya, N., Jingami, H., and Morikawa, K. Structural views of the ligand-binding cores of a metabotropic glutamate receptor complexed with an antagonist and both glutamate and Gd3+. *Proceedings of the National Academy of Sciences*, 99(5):2660–2665, 2002.

[101] Muto, T., Tsuchiya, D., Morikawa, K., and Jingami, H. Structures of the extracellular regions of the group II/III metabotropic glutamate receptors. *Proceedings of the National Academy of Sciences*, 104(10):3759–3764, 2007.

[102] Monn, J. A., Prieto, L., Taboada, L., Hao, J., Reinhard, M. R., Henry, S. S., Beadle, C. D., Walton, L., Man, T., Rudyk, H., et al. Synthesis and pharmacological characterization of C4-(Thiotriazolyl)-substituted-2-aminobicyclo [3.1.0] hexane-2, 6-dicarboxylates. identification of (1R, 2S, 4R, 5R, 6R)-2-amino-4-(1H-1, 2, 4-triazol-3-ylsulfanyl) bicyclo [3.1.0] hexane-2, 6-dicarboxylic acid (LY2812223), a highly potent, functionally selective mGlu2 receptor agonist. *Journal of medicinal chemistry*, 58(18):7526–7548, 2015.

[103] Chappell, M. D., Li, R., Smith, S. C., Dressman, B. A., Tromiczak, E. G., Tripp, A. E., Blanco, M.-J., Vetman, T., Quimby, S. J., Matt, J., et al. Discovery of (1S, 2R, 3S, 4S, 5R, 6R)-2-amino-3-[(3, 4-difluorophenyl) sulfanylmethyl]-4-hydroxy-bicyclo [3.1.0] hexane-2, 6-dicarboxylic acid hydrochloride (LY3020371·HCl): A potent, metabotropic glutamate 2/3 receptor antagonist with

antidepressant-like activity. *Journal of medicinal chemistry*, 59(24):10974–10993, 2016.

[104] Monn, J. A., Henry, S. S., Massey, S. M., Clawson, D. K., Chen, Q., Diseroad, B. A., Bhardwaj, R. M., Atwell, S., Lu, F., Wang, J., et al. Synthesis and pharmacological characterization of $C4_{\beta}$-amide-substituted 2-aminobicyclo [3.1.0] hexane-2, 6-dicarboxylates. identification of (1S, 2S, 4S, 5R, 6S)-2-Amino-4-[(3-methoxybenzoyl) amino] bicyclo [3.1.0] hexane-2, 6-dicarboxylic acid (LY2794193), a highly potent and selective mGlu3 receptor agonist. *Journal of medicinal chemistry*, 61(6):2303–2328, 2018.

[105] Silbering, A. F. and Benton, R. Ionotropic and metabotropic mechanisms in chemoreception:'chance or design'? *EMBO reports*, 11(3):173–179, 2010.

[106] Rimal, S. and Lee, Y. The multidimensional ionotropic receptors of Drosophila melanogaster. *Insect molecular biology*, 27(1):1–7, 2018.

[107] Clyne, P. J., Warr, C. G., Freeman, M. R., Lessing, D., Kim, J., and Carlson, J. R. A novel family of divergent seven-transmembrane proteins: candidate odorant receptors in Drosophila. *Neuron*, 22(2):327–338, 1999.

[108] Joseph, R. M. and Carlson, J. R. Drosophila chemoreceptors: a molecular interface between the chemical world and the brain. *Trends in Genetics*, 31(12):683–695, 2015.

[109] Benton, R., Vannice, K. S., Gomez-Diaz, C., and Vosshall, L. B. Variant ionotropic glutamate receptors as chemosensory receptors in Drosophila. *Cell*, 136(1):149–162, 2009.

[110] Silbering, A. F., Rytz, R., Grosjean, Y., Abuin, L., Ramdya, P., Jefferis, G. S., and Benton, R. Complementary function and integrated wiring of the evolutionarily distinct Drosophila olfactory subsystems. *Journal of Neuroscience*, 31(38):13357–13375, 2011.

[111] Sánchez-Alcañiz, J. A., Silbering, A. F., Croset, V., Zappia, G., Sivasubramaniam, A. K., Abuin, L., Sahai, S. Y., Auer, T. O., Cruchet, S., Neagu-Maier, G. L., et al. An expression atlas of chemosensory ionotropic glutamate receptors identifies a molecular basis of carbonation detection. Technical report, Université de Fribourg, 2018.

[112] Prieto-Godino, L. L., Rytz, R., Cruchet, S., Bargeton, B., Abuin, L., Silbering, A. F., Ruta, V., Dal Peraro, M., and Benton, R. Evolution of acid-sensing olfactory circuits in drosophilids. *Neuron*, 93(3):661–676, 2017.

[113] Croset, V., Schleyer, M., Arguello, J. R., Gerber, B., and Benton, R. A molecular and neuronal basis for amino acid sensing in the Drosophila larva. *Scientific reports*, 6:34871, 2016.

[114] Rytz, R., Croset, V., and Benton, R. Ionotropic receptors (IRs): chemosensory ionotropic glutamate receptors in Drosophila and beyond. *Insect biochemistry and molecular biology*, 43(9):888–897, 2013.

[115] Ni, L., Klein, M., Svec, K. V., Budelli, G., Chang, E. C., Ferrer, A. J., Benton, R., Samuel, A. D., and Garrity, P. A. The ionotropic receptors IR21a and IR25a mediate cool sensing in Drosophila. *Elife*, 5:e13254, 2016.

[116] Chen, C., Buhl, E., Xu, M., Croset, V., Rees, J. S., Lilley, K. S., Benton, R., Hodge, J. J., and Stanewsky, R. Drosophila ionotropic receptor 25a mediates circadian clock resetting by temperature. *Nature*, 527(7579):516, 2015.

[117] Knecht, Z. A., Silbering, A. F., Ni, L., Klein, M., Budelli, G., Bell, R., Abuin, L., Ferrer, A. J., Samuel, A. D., Benton, R., et al. Distinct combinations of variant ionotropic glutamate receptors mediate thermosensation and hygrosensation in Drosophila. *Elife*, 5:e17879, 2016.

[118] Knecht, Z. A., Silbering, A. F., Cruz, J., Yang, L., Croset, V., Benton, R., and Garrity, P. A. Ionotropic receptor-dependent moist and dry cells control hygrosensation in Drosophila. *Elife*, 6:e26654, 2017.

[119] Koh, T.-W., He, Z., Gorur-Shandilya, S., Menuz, K., Larter, N. K., Stewart, S., and Carlson, J. R. The Drosophila IR20a clade of ionotropic receptors are candidate taste and pheromone receptors. *Neuron*, 83(4):850–865, 2014.

[120] Silbering, A. F., Bell, R., Münch, D., Cruchet, S., Gomez-Diaz, C., Laudes, T., Galizia, C. G., and Benton, R. IR40a neurons are not DEET detectors. *Nature*, 534(7608):E5, 2016.

[121] Abuin, L., Bargeton, B., Ulbrich, M. H., Isacoff, E. Y., Kellenberger, S., and Benton, R. Functional architecture of olfactory ionotropic glutamate receptors. *Neuron*, 69(1):44–60, 2011.

Chapter 2 - Computing Conformational Free Energies of iGluR Ligand-Binding Domains

Reproduced in part from:
Ionotropic Glutamate Receptor Technologies, Neuromethods vol. 106 chap 9
DOI 10.1007/978-1-4939-2812-5_9, Springer Science+Business Media New York
2016

by

Alvin Yu, Tyler Wied, John Belcher, and Albert Y. Lau

Ionotropic glutamate receptors (iGluRs) transduce chemical signals at synapses into electrical impulses. This function relies on concerted conformational changes that are propagated among the linked domains of the tetrameric protein assembly making up each receptor. A key conformational change is the closure of the ligand-binding domain (LBD) upon agonist binding, which eventually gates the transmembrane ion channel domain. The free energy that becomes available for gating transitions is governed by the LBD free energy landscapes for apo and ligand-bound states. These landscapes describe the thermodynamic equilibrium among various LBD conformations. Delineating these landscapes is essential for understanding the molecular driving forces underlying iGluR function. Molecular dynamics free energy simulations offer a means for estimating these quantities, which are difficult to extract from experimental results alone. Here, we describe the process of carrying out a free energy computation using an umbrella sampling strategy for characterizing large-scale conformational changes in iGluR LBDs.

2.1 Background

Structural studies of ionotropic glutamate receptors (iGluRs) have provided valuable insights into the molecular mechanics of activation and desensitization [1]. These studies have helped characterize the conformational changes that take place within an isolated ligand-binding domain (LBD) upon binding of agonists and antagonists in addition to the molecular rearrangements that are associated with desensitization. The numerous crystal structures of isolated iGluR domains and intact iGluR receptors also provide atomistic models that enable molecular simulation studies. Such studies allow one to probe conformational energetics not easily accessed by experimental approaches alone by, for example, considering conformational states that are transient, not heavily populated, or refractory to crystallization. Such considerations are important for understanding iGluRs, which are allosteric receptors that rely on complex molecular dynamics for executing their function.

Straightforward, unbiased molecular dynamics (MD) simulations have shed light on important dynamic events involving the LBD; for example, by suggesting possible pathways for conformational changes related to receptor activation [2, 3] and rearrangements in the configuration of waters in the binding cleft [4]. While these types of simulations are useful, they can have limitations when applied to complex macromolecular systems that exhibit significant conformational changes. Instead, one can employ a simulation strategy that generates a "free energy landscape", or equivalently, a "potential of mean force" (PMF), along specifically chosen coordinates to obtain a reduced description of conformational changes in the system. The free energy difference between conformational states reflects the probability of finding the system in those states.

This chapter focuses on the computation of free energy landscapes using umbrella sampling MD simulations. The systems of interest here are iGluR LBDs. Due to space constraints, we limit the discussion of numerous general aspects of setting up and carrying out an MD simulation. These topics, however, are covered elsewhere in excellent overviews, e.g., by Cheng and Ivanov [5] and the NAMD tutorials at www.ks.uiuc.edu/Training/Tutorials/namd-index.html. The theory of umbrella sampling is discussed in some detail for the sake of completeness and because of its practical utility, as it provides a guide for writing the computer code necessary for carrying out the calculations. We hope that this chapter will be useful to readers in facilitating computational studies of iGluR conformational energetics.

2.2 Materials

The most widely used molecular dynamics (MD) simulation packages for biomolecules are CHARMM [6], AMBER [7], GROMOS [8], NAMD [9], and GROMACS [10]. CHARMM and AMBER have been in continuous development for the longest, offer advanced sampling and analysis capabilities, and are associated with their own force fields for proteins, nucleic acids, lipids, and carbohydrates [5]. All LBD umbrella sampling simulations described below were performed using CHARMM, although other packages can be used as well. NAMD and GROMACS have been developed to excel in performance (i.e., computational speed), and they support numerous force fields. The package Desmond [11], from D.E. Shaw Research, also offers high performance and supports multiple force fields. For a comparison of different force fields, see a review by Guvench and MacKerell [12]. VMD [13] is a widely used molecular graphics program for structure and trajectory analysis.

MD packages generally require an operating system based on Unix/Linux. Some aspects of setting up a simulation system may be accomplished on a workstation, but equilibration and production simulations are typically performed on either commodity computing clusters containing many processors or high-end computing platforms available through resources such as the Extreme Science and Engineering Discovery Environment (XSEDE; www.xsede.org).

2.3 Methods

2.3.1 Molecular Dynamics Simulations

At the heart of a molecular mechanics simulation is the potential energy function that relates the atomic coordinates of a system, $\mathbf{R}$, to the energy, U [14] (**Fig. 2-1**). The minimal set of energy terms commonly used in biomolecular simulations is the following:

$$\begin{aligned} U(\mathbf{R}) = \sum_{\text{bonds}} K_b(b-b_0)^2 + \sum_{\text{angles}} K_\theta(\theta-\theta_0)^2 + \\ \sum_{\text{dihedrals}} K_\phi[1+\cos(n\phi-\delta)] + \sum_{\substack{\text{impropers}\\ \text{dihedrals}}} K_\chi(\chi-\chi_0)^2 + \\ \sum_{\text{bonded}} \left\{ \frac{q_i q_j}{\epsilon r_{ij}} + \epsilon_{ij}\left[\left(\frac{R_{\min,ij}}{r_{ij}}\right)^{12} - 2\left(\frac{R_{\min,ij}}{r_{ij}}\right)^{6}\right]\right\} \end{aligned} \tag{2.1}$$

where the equilibrium values for the bond distance, b, the valence angle, θ, and the improper dihedral angle (to maintain chiral and planar centers), χ, are b_0, θ_0, and χ_0, respectively; the multiplicity and phase for the dihedral angle, ϕ, are n and δ; q_i and q_j are the partial atomic charges on atoms i and j; the Ks are the force constants; ϵ is the dielectric constant; r_{ij} is the distance between atoms i and j; and the Lennard-Jones potential well depth and minimum interaction radius are ϵ_{ij} and $R_{\min,ij}$. The "bonded" terms are the bonds, angles, dihedrals, and improper dihedrals; the "non-bonded" terms are the electrostatic interactions treated using a Coulombic term and the interatomic repulsion and dispersion interactions treated using a Lennard-Jones "6-12" term. These terms are common to most modern biomolecular force fields, including CHARMM, AMBER, GROMOS, and OPLS [15], among others. To increase the accuracy of force fields, additional or alternate energy terms have been incorporated. An example of an additional term is the two-dimensional dihedral energy grid "correction map" (CMAP) applied to protein backbone ϕ and ψ angles in CHARMM [16]. Other examples are described in a review by MacKerell [14]. Once $U(\mathbf{R})$ is defined, the time evolution of a system of atoms may be obtained by integrating Newton's equations of motion,

$$F(\mathbf{R}(t)) = -\nabla U(\mathbf{R}(t)) = M\ddot{\mathbf{R}}(t) \tag{2.2}$$

where t is time, F is the ensemble of forces on the atoms, ∇U is the gradient of the potential, M is the matrix that specifies the mass of every atom, and $\ddot{\mathbf{R}}$ is the second derivative of the atomic coordinates with respect to time. Combining Taylor series expansions for $\mathbf{R}(t+\delta t)$ and $\mathbf{R}(t-\delta t)$ yield the Verlet integration algorithm, which is used to generate the dynamical trajectory of the system. Details concerning the propagation of atomic velocities and positions, as well as many practical considerations regarding MD simulations, are covered in a review by Cheng and Ivanov [5].

2.3.2 Order Parameters

How might one go about computing the change in free energy for a conformational transition in a system? A first step would be to identify a continuous coordinate ξ, also referred to as an order parameter, which is a function of a few or more degrees of freedom in the system, and traverses the system's conformational states of interest. Order parameters along the most important degrees of freedom are often chosen as a reduced descriptor of the system. Any choice of order parameter is possible, such as a distance, an angle, or a root mean square deviation from a reference state. The order parameter may be multidimensional, e.g., two distances or a distance plus an angle. Choosing an appropriate order parameter for a given system often requires a mix of physical intuition and prior knowledge of the chemical process under consideration. Several novel strategies have been developed to quantitatively determine order parameters [17]. These techniques include combining MD simulations with either principal component analysis (PCA) or a determination of the atomic fluctuations that correspond to the largest amounts of information transfer in a conformational transition. Normal mode analysis (NMA) has also been suggested as a means for identifying the most important collective motions [18–20].

2.3.3 Umbrella Sampling

Once the order parameter, ξ, is chosen, the average distribution function along ξ can be provided by a Boltzmann-weighted average [21, 22]:

$$\langle \rho(\xi) \rangle = \int \frac{\delta[\xi'(\mathbf{r}) - \xi] e^{-U(\mathbf{r})/k_B T} d\mathbf{r}}{\int e^{-U(\mathbf{r})/k_B\ T} d\mathbf{r}} \tag{2.3}$$

where $U(\mathbf{r})$ is the energy of the system as a function of the atomic coordinates $\mathbf{r}$, $\delta[\cdot]$ is the Dirac delta function, k_B is Boltzmann's constant, and T is temperature. The free energy, or PMF, $W(\xi)$, along ξ can be computed as follows:

$$W(\xi) - W(\xi') = -k_B T \ln \left[\frac{\langle \rho(\xi) \rangle}{\langle \rho(\xi') \rangle} \right] \tag{2.4}$$

where ξ' is an arbitrary reference value. In principle, $W(\xi)$ could be computed directly from a "brute force" molecular dynamics simulation of the system. In practice, however, this approach is often limited because large free energy barriers along ξ can impede the amount of conformational space that is sampled within an allotted simulation time (**Fig. 2-2**). Poor sampling would yield unreliable statistics from which to calculate free energy differences. Consequently, numerous sampling strategies have been developed to more effectively calculate a PMF from molecular dynamics simulation trajectories. One of these strategies is the umbrella sampling method developed by Torrie and Valleau [23]. This method typically involves carrying out multiple independent simulations, where each simulation takes place in the presence of an applied potential, $w(\xi)$. This artificial bias for each simulation, or window, confines sampling to a region near a chosen value of ξ. The chosen ξ is different for each window. A convenient form for the artificial bias in the ith window is a harmonic function, i.e., $w_i = \frac{1}{2}K(\xi - \xi_i)^2$, where K is the force constant. The set of ξ_i spans the conformational space of interest.

A biased distribution function is generated by each window. The process of unbiasing and recombining the results of sampling in all the windows in order to obtain $W(\xi)$ involves calculating the free energy associated with introducing the window potential. Numerous approaches have been developed to accomplish this task [21–24]. We will discuss only one of these approaches, the weighted histogram analysis method (WHAM) [25], although other valid approaches are available. Essentially, the WHAM estimates the unbiased distribution function by computing a weighted sum of all the sampling data and determining the functional form of the weight factors that minimizes the statistical error [21, 25, 26].

Briefly, WHAM uses two equations, known as the WHAM equations, to compute the unbiased distribution and potential of mean force. The first equation is for the unbiased distribution function, $\langle\rho(\xi)\rangle$, and is given by

$$\langle\rho(\xi)\rangle = \sum_{i=1}^{N} n_i \ \langle\rho(\xi)\rangle_{(i)} \times \left[\sum_{j=1}^{N} n_j \ e^{-[w_j(\xi)-F_j]/k_BT}\right]^{-1} \tag{2.5}$$

where $[\langle\rho(\xi)\rangle]_{(i),\text{unbiased}}$ are the individual unbiased distribution functions, and n_i are the number of sampling data points used to generate each biased distribution function.

second equation gives F_i, given by,

$$e^{-F_i/k_BT} = \int \langle\rho(\xi)\rangle e^{-w_i(\xi)/k_BT} d\xi \tag{2.6}$$

These two equations are interdependent (i.e., 2.5 requires $\langle\rho(\xi)\rangle$ and 2.6 requires F_i) and are therefore solved in a self-consistent manner via an iterative procedure. First, an estimate for the unbiased distribution $\langle\rho(\xi)\rangle$ is computed using an initial guess for all of the F_j. Next, $e^{-F_i(\xi)/k_BT}$ is evaluated using the estimate for $\langle\rho(\xi)\rangle$. The cycle is repeated using the new estimates for F_j until both equations have converged. The equations above are written in terms of a one-dimensional ξ. The approach can be applied to multidimensional scenarios as well. For example, the WHAM equations can be reformulated for handling a two-dimensional order parameter involving the variables ξ_1 and ξ_2.

Additional theoretical details and derivations may be found in Souaille and Roux [27]. Generic computer code for carrying out WHAM calculations is also provided in this reference. The first PMFs computed for an iGluR LBD involved the GluA2 S1S2 construct [28]. The umbrella sampling simulations generating the PMFs employed a two-dimensional order parameter consisting of two distances between the lobes of the LBD (**Fig. 2-3**). The order parameter was inspired by the observation that the distances between T480 and S654, and between E402 and T686 each differ by ∼3.5 Å when comparing crystal structures of apo and agonist-bound LBDs [29]. Each pair of residues is hydrogen bonded in the agonist-bound structures but not in the apo structure. In the order parameter (ξ_1, ξ_2), ξ_1 is the distance between the centers of mass (COM) of residues 479-481 in lobe 1 and residues 654-655 in lobe 2, and ξ_2 is the distance between the COM of residues 401-403 in lobe 1 and residues 686-687 in lobe 2. These residues correspond to regions that are not in flexible loops. All non-hydrogen atoms were included in the selections. The rationale for using multiple residues in the definitions of ξ_1 and ξ_2 was to increase the likelihood that changes in the COM distances would represent large-scale conformational changes rather than fluctuations in individual residues. We should note that we initially attempted to define an angle to monitor opening and closing of the LBD, in a fashion similar to how the crystal structures were characterized. However, we found this order parameter behaved poorly and is ill-suited for use with the umbrella sampling restraining potentials. Other order parameters have been used by other investigators, for example, the distance between the Cα atoms of G451 and S652 [30]. Our chosen order parameter was intended to sufficiently capture the principal conformational transition inferred from LBD crystal structures. We acknowledge, however, that an order parameter with higher dimensionality would be required to capture motions involving additional degrees of freedom.

In our previous studies [28, 31, 32], we have found that umbrella sampling windows spaced 1 Å apart along ξ_1 and ξ_2 provide sufficient sampling overlap between neighboring windows. Windows spanning $(\xi_1, \xi_2) = (8, 6)$ to (18,17) (in Å) appear to provide reasonable coverage of conformational space (**Fig. 2-4**). Initially, for GluA2, ~200 windows were used [28]. Many of these windows, however, turned out to reside in regions of high conformational free energy. Subsequently, for GluN1, GluN2A, and GluN3A, the number of windows was cut to the ~100 that contribute most to $\langle \rho(\xi_1, \xi_2) \rangle$ [32]. Windows in the "corner" regions, which correspond to either large ξ_1 and small ξ_2 or small ξ_1 and large ξ_2, are the windows that are generally very high in conformational free energy and can be safely omitted from the computation.

2.3.4 Setting Up the System

Molecular simulations require complete protein models; that is, any missing non-hydrogen side chain or backbone atoms in a PDB entry must be built into the model. For building missing sidechains, we routinely use the program SCWRL [33]. For building missing backbone atoms, we use the loop-modeling routine of the program MODELER [34], also available via the ModLoop server (modbase.compbio.ucsf.edu). The next task is to generate initial protein configurations for each of the umbrella sampling windows. We use a targeted (biased potential) MD procedure to generate these LBD configurations. In this procedure, the Cα atoms of lobes 1 and 2 are separately restrained to their crystal structure configuration using an RMSD restraining potential while the lobes are pushed apart by performing dynamics in the presence of biasing potentials applied separately to ξ_1 and ξ_2. These simulations are run in vacuum. The intent of the RMSD restraints is to prevent the lobes from partially unfolding while ξ_1 and ξ_2 are rapidly moved apart. First, ξ_1 and ξ_2 are simultaneously incremented starting from a closed-conformation crystal structure. Next, ξ_1 and ξ_2 are individually incremented starting from configurations generated in the previous step. System setup concludes with solvating each conformer with explicit water and 100-150 mM NaCl. Nonzero total charge on the protein (and ligand, if present) may be neutralized by using excess Na^+ or Cl^- ions. We retain the positions of crystallographically-determined water molecules within the binding cleft.

2.3.5 Small Molecule Parametrization

For iGluR ligands (small organic molecules) that are not amino acids, the molecular mechanics parameters and charges will need to be determined. This can be accomplished a number of ways. One approach is to use the general AMBER force field (GAFF) [35] together with Antechamber, a toolkit for automatic atom typing [36]. For CHARMM-consistent parameters, the CHARMM general force field (CGenFF) program can be used for automatic atom typing and assignment of parameters and charges by analogy [37,38]. Alternatively, a method called general automated atomic model parameterization (GAAMP) can be used to generate parameters based on the results of *ab initio* quantum mechanical (QM) calculations [39]. GAAMP can also be used to optimize existing parameter sets. Two web portals are available for automatic parameterization of small compounds using CGenFF: ParamChem (www.paramchem.org) and MATCH (brooks.chem.lsa.umich.edu). A web portal for GAAMP is at gaamp.lcrc.anl.gov. Finally, parameterization can be performed manually. Our approach for DNQX [28] included the following steps: (1) charge fit a geometry-optimized molecule to QM electrostatic potential maps, (2) optimize force constants by reproducing vibrational frequencies and potential energy distributions from QM calculations, and (3) optimize dihedral angle parameters by reproducing dihedral potential energy surfaces from QM calculations.

2.3.6 Equilibriation and Production

Each system must be equilibrated before sampling data are collected. The steps we routinely take are described below:

- First, the protein and ligand atoms are held fixed while dynamics are carried out on the solvent.
- Next, dynamics are carried out in the presence of stabilizing restraining potentials on the protein and ligand. These restraints are gradually reduced over the course

of the equilibration, which is carried out in the constant atom number, volume, and temperature (NVT) ensemble.

- Finally, after these restraints have been removed, the simulation is transitioned to a constant pressure (NPT) ensemble.

In all steps of the equilibration involving non-fixed protein and ligand atoms, the umbrella restraints on ξ_1 and ξ_2 are also applied, as well as weak restraints on the COM of three regions in the core of lobe 1. The latter restraint, which does not hinder inter-lobe dynamics, prevents translation and rotation of the protein during the production phase of the simulations.

The most effective way to implement the restraints described above vary depending on which MD simulation package is used. For example, in CHARMM, we use the restraint invoked by the keywords "CONS HARM" for the stabilizing potential. The "MMFP GEO SPHERE RCM DISTANCE" restraint is used for ξ_1 and ξ_2, and the "CONS HMCM" restraint is used for translation/rotation. The "MMFP" module allows the instantaneous values of ξ_1 and ξ_2 to be written on-the-fly. Details specifying how to use these keywords are included in CHARMM's "doc" (documentation) directory.

The simulations enter production phase after equilibration has completed. Here, the values for ξ_1 and ξ_2 in each window i are recorded, which contribute to the biased distribution function $\langle\rho(\xi_1,\xi_2)\rangle_{(i)}$. It is critical that all distributions are carefully monitored to make sure there is sufficient overlap between neighboring windows, which is a requirement of WHAM [24]. If deficiencies are observed, additional windows at intermediate positions may be added. Additional windows may also be added to expand the boundaries of (ξ_1,ξ_2) if, after computing $W(\xi_1,\xi_2)$, it is determined that the free energy basins have not been sufficiently mapped. Automated adaptive umbrella sampling approaches may also be useful for optimizing the task of selecting windows [40].

2.3.7 Error Analysis

Once $W(\xi_1, \xi_2)$ has been computed, the quality of the sampling can be assessed. Two methods of analysis are block averaging and bootstrapping. In block averaging, as the name suggests, the sampling data (e.g., (ξ_1, ξ_2)) are divided into M segments, or "blocks", corresponding to a given simulation time. The PMF is then calculated for each block, and error bars can be generated to represent the deviation observed in the M PMFs. For a recent example of block averaging applied to umbrella sampling, see Park et al. [41]. In bootstrapping [42], "resampled" data sets are constructed by randomly selecting data points from the original sampling data. The resampling takes place using time intervals estimated to correspond to the time it takes for correlated fluctuations in (ξ_1, ξ_2) to decay to zero. Previously, we used time intervals of 300 ps for GluN LBDs and 200 ps for the GluA2 LBD [32]. Resampling is repeated until each resampled data set is the same size as the original. The PMF is then calculated for each resampled set, and error bars can be generated as described above.

Ensuring that convergence in the sampling has been reached, using, for example, one or both of the methods described above, is an important step in molecular simulation studies. It is also a prerequisite for reasonably comparing the computed free energies with experimental measurements. For the computations described in this chapter, which are focused on large-scale conformational transitions, experimental measurements that, at least in principle, can be compared fairly directly include those that characterize conformational ensembles in solution, such as small-angle X-ray scattering (SAXS) [28, 43], and nuclear magnetic resonance (NMR) [44] and fluorescence resonance energy transfer (FRET) spectroscopy [45]. The computed free energy landscapes provide Boltzmann weights to be applied to the molecular configurations extracted from the simulations, which in turn provide a framework, based on molecular ensembles, for interpreting the measured data.

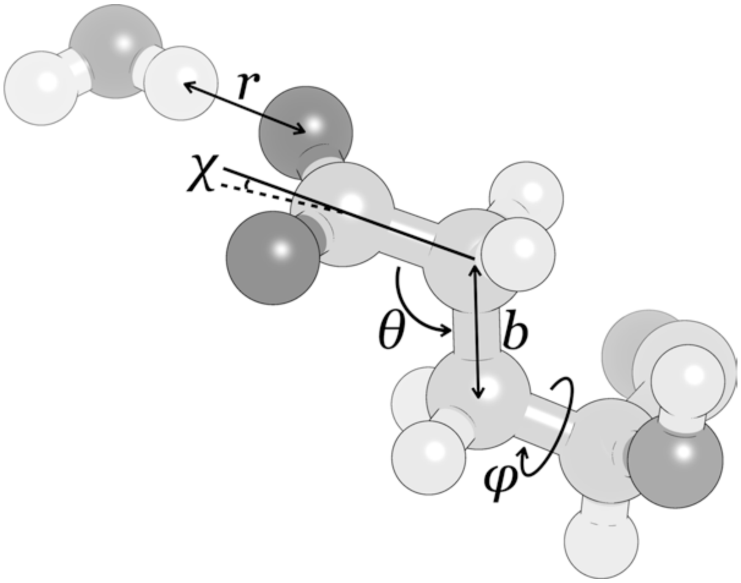


Fig. 2-1: **A molecular mechanics potential energy function** typically includes terms describing bond stretching (b), bond angle bending (θ), bond twisting (ϕ), improper dihedrals (χ), and non-bonded interactions (r).

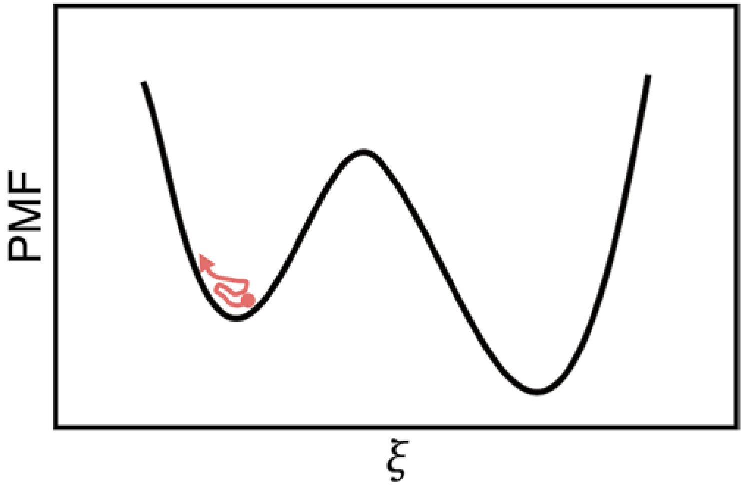


Fig. 2-2: **Free energy barriers.** In unbiased molecular simulations, one or more large free energy barriers along an order parameter, ξ, can impede conformational sampling (depicted in *red*).

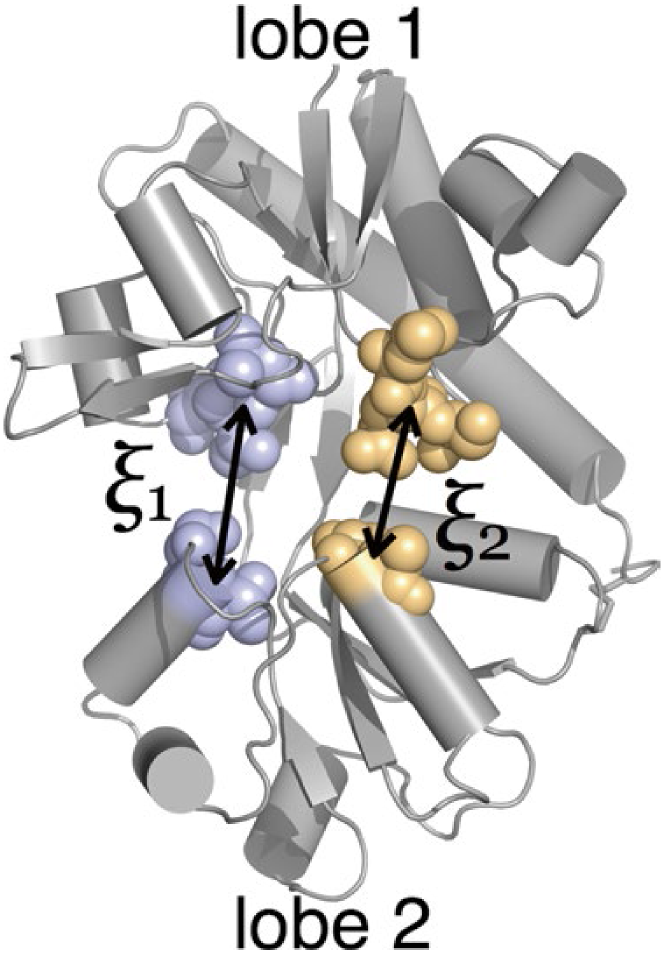


Fig. 2-3: **Order parameters.** Our chosen order parameter, (ξ_1, ξ_2), used to describe large-scale conformational transitions in the GluA2 LBD. ξ_1 and ξ_2 indicate the distances between the centers of mass of the atoms that are shown as spheres. In this figure, (ξ_1, ξ_2) = (12.8 Å, 11.4 Å)

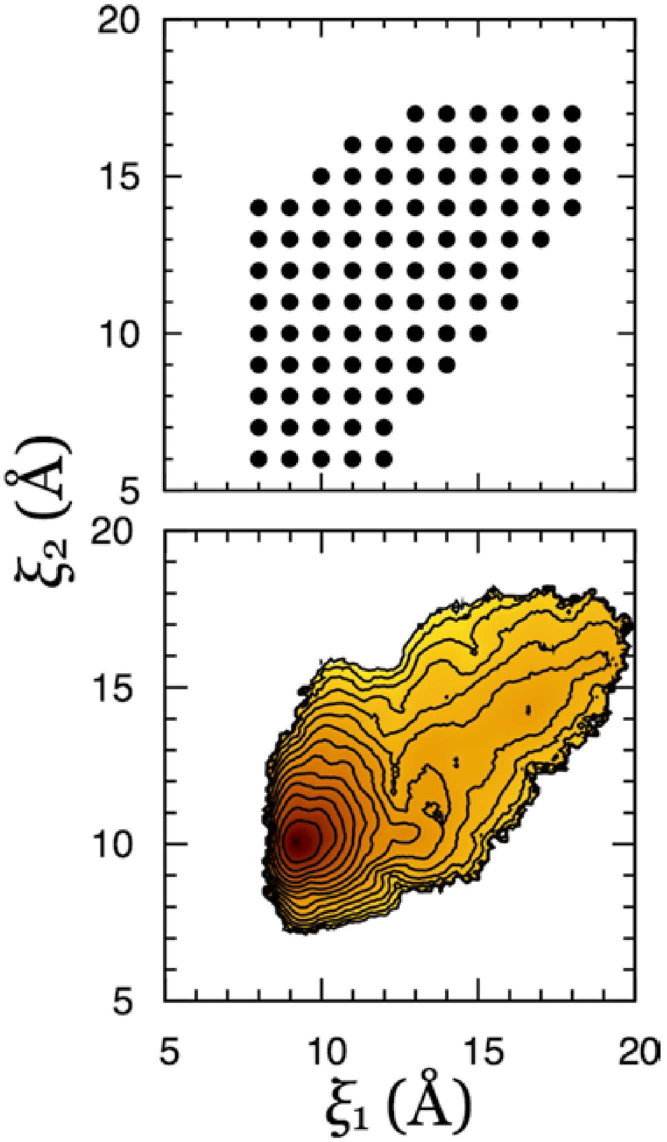


Fig. 2-4: **Umbrella sampling** involving the GluN1 glycine-bound LBD. The top panel shows the umbrella restraining potential equilibrium positions of the 93 windows that produce the free energy landscape shown in the bottom panel. In the bottom panel, each contour level corresponds to a difference of 1 kcal/mol, and darker regions are lower in free energy. See Yao et al. [32] for details.

References

[1] Kumar, J. and Mayer, M. L. Functional Insights from Glutamate Receptor Ion Channel Structures. *Annual review of physiology*, 75:313–337, 2013.

[2] Arinaminpathy, Y., Sansom, M. S. P., and Biggin, P. C. Binding site flexibility: molecular simulation of partial and full agonists within a glutamate receptor. *Molecular Pharmacology*, 69(1):11–18, January 2006.

[3] Mamonova, T., Speranskiy, K., and Kurnikova, M. Interplay between structural rigidity and electrostatic interactions in the ligand binding domain of GluR2. *Proteins: Structure, Function, and Bioinformatics*, 73(3):656–671, November 2008.

[4] Vijayan, R., Sahai, M. A., Czajkowski, T., and Biggin, P. C. A comparative analysis of the role of water in the binding pockets of ionotropic glutamate receptors. *Physical Chemistry Chemical Physics*, 12(42):14057–14066, October 2010.

[5] Cheng, X. and Ivanov, I. Molecular Dynamics. In *Computational Toxicology*, Methods in Molecular Biology, pages 243–285. Humana Press, Totowa, NJ, 2012. DOI: 10.1007/978-1-62703-050-2_11.

[6] Brooks, B., Brooks, C., MacKerell, A., Nilsson, L., Petrella, R., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R., Post, C., Pu, J., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D., and Karplus, M. CHARMM: The Biomolecular Simulation Program. *Journal of computational chemistry*, 30(10):1545–1614, July 2009.

[7] Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic

Molecules. *Journal of the American Chemical Society*, 117(19):5179–5197, May 1995.

[8] Scott, W. R. P., Hunenberger, P. H., Tironi, I. G., Mark, A. E., Billeter, S. R., Fennen, J., Torda, A. E., Huber, T., Kruger, P., and van Gunsteren, W. F. The GROMOS Biomolecular Simulation Program Package. *The Journal of Physical Chemistry A*, 103(19):3596–3607, May 1999.

[9] Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, 26(16):1781–1802, December 2005.

[10] Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., and Berendsen, H. J. C. GROMACS: Fast, flexible, and free. *Journal of Computational Chemistry*, 26(16):1701–1718, December 2005.

[11] Bowers, K. J., Chow, E., Xu, H., Dror, R. O., Eastwood, M. P., Gregersen, B. A., Klepeis, J. L., Kolossvary, I., Moraes, M. A., Sacerdoti, F. D., Salmon, J. K., Shan, Y., and Shaw, D. E. Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. In *Proceedings of the 2006 ACM/IEEE Conference on Supercomputing*, SC '06, New York, NY, USA, 2006. ACM.

[12] Guvench, O. and MacKerell, A. D. Comparison of Protein Force Fields for Molecular Dynamics Simulations. In *Molecular Modeling of Proteins*, Methods Molecular Biology, pages 63–88. Humana Press, 2008. DOI: 10.1007/978-1-59745-177-2_4.

[13] Humphrey, W., Dalke, A., and Schulten, K. VMD: Visual molecular dynamics. *Journal of Molecular Graphics*, 14(1):33–38, February 1996.

[14] Mackerell, A. D. Empirical force fields for biological macromolecules: Overview and issues. *Journal of Computational Chemistry*, 25(13):1584–1604, October 2004.

[15] Jorgensen, W. L. and Tirado-Rives, J. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society*, 110(6):1657–1666, March 1988.

[16] MacKerell, A. D., Feig, M., and Brooks, C. L. Improved Treatment of the Protein Backbone in Empirical Force Fields. *Journal of the American Chemical Society*, 126(3):698–699, January 2004.

[17] Perilla, J. R. and Woolf, T. B. Towards the prediction of order parameters from molecular dynamics simulations in proteins. *The Journal of Chemical Physics*, 136(16), April 2012.

[18] Alexandrov, V., Lehnert, U., Echols, N., Milburn, D., Engelman, D., and Gerstein, M. Normal modes for predicting protein motions: A comprehensive database assessment and associated Web tool. *Protein Science : A Publication of the Protein Society*, 14(3):633–643, March 2005.

[19] Chennubhotla, C. and Bahar, I. Signal propagation in proteins and relation to equilibrium fluctuations. *PLoS computational biology*, 3(9):1716–1726, September 2007.

[20] Skjaerven, L., Hollup, S. M., and Reuter, N. Normal mode analysis for proteins. *Journal of Molecular Structure: THEOCHEM*, 898(1):42–48, March 2009.

[21] Kirkwood, J. G. Statistical Mechanics of Fluid Mixtures. *The Journal of Chemical Physics*, 3(5):300–313, May 1935.

[22] Roux, B. The calculation of the potential of mean force using computer simulations. *Computer Physics Communications*, 91(1):275–282, September 1995.

[23] Torrie, G. M. and Valleau, J. P. Monte Carlo free energy estimates using non-Boltzmann sampling: Application to the sub-critical Lennard-Jones fluid. *Chemical Physics Letters*, 28(4):578–581, October 1974.

[24] Kastner, J. Umbrella sampling. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 1(6):932–942, November 2011.

[25] Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., and Kollman, P. A. THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *Journal of Computational Chemistry*, 13(8):1011–1021, October 1992.

[26] Ferrenberg, A. M. and Swendsen, R. H. Optimized Monte Carlo data analysis. *Physical Review Letters*, 63(12):1195–1198, September 1989.

[27] Souaille, M. and Roux, B. Extension to the weighted histogram analysis method: combining umbrella sampling with free energy calculations. *Computer Physics Communications*, 135(1):40–57, March 2001.

[28] Lau, A. Y. and Roux, B. The Free Energy Landscapes Governing Conformational Changes in a Glutamate Receptor Ligand-Binding Domain. *Structure (London, England : 1993)*, 15(10):1203–1214, October 2007.

[29] Armstrong, N. and Gouaux, E. Mechanisms for activation and antagonism of an AMPA-sensitive glutamate receptor: crystal structures of the GluR2 ligand binding core. *Neuron*, 28(1):165–181, October 2000.

[30] Mamonova, T., Yonkunas, M. J., and Kurnikova, M. G. Energetics of the cleft closing transition and the role of electrostatic interactions in conformational rearrangements of the glutamate receptor ligand binding domain. *Biochemistry*, 47(42):11077–11085, October 2008.

[31] Lau, A. Y. and Roux, B. The hidden energetics of ligand binding and activation in a glutamate receptor. *Nature Structural & Molecular Biology*, 18(3):283–287, March 2011.

[32] Yao, Y., Belcher, J., Berger, A. J., Mayer, M. L., and Lau, A. Y. Conformational analysis of NMDA receptor GluN1, GluN2, and GluN3 ligand-binding

domains reveals subtype-specific characteristics. *Structure (London, England: 1993)*, 21(10):1788–1799, October 2013.

[33] Krivov, G. G., Shapovalov, M. V., and Dunbrack, R. L. Improved prediction of protein side-chain conformations with SCWRL4. *Proteins*, 77(4):778–795, December 2009.

[34] Fiser, A. and Sali, A. ModLoop: automated modeling of loops in protein structures. *Bioinformatics (Oxford, England)*, 19(18):2500–2501, December 2003.

[35] Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. Development and testing of a general amber force field. *Journal of Computational Chemistry*, 25(9):1157–1174, July 2004.

[36] Wang, J., Wang, W., Kollman, P. A., and Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *Journal of Molecular Graphics and Modelling*, 25(2):247–260, October 2006.

[37] Vanommeslaeghe, K. and MacKerell, A. D. Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing. *Journal of chemical information and modeling*, 52(12):3144–3154, December 2012.

[38] Vanommeslaeghe, K., Raman, E. P., and MacKerell, A. D. Automation of the CHARMM General Force Field (CGenFF) II: Assignment of bonded parameters and partial atomic charges. *Journal of chemical information and modeling*, 52(12):3155–3168, December 2012.

[39] Huang, L. and Roux, B. Automated Force Field Parameterization for Nonpolarizable and Polarizable Atomic Models Based on Ab Initio Target Data. *Journal of Chemical Theory and Computation*, 9(8):3543–3556, August 2013.

[40] Wojtas-Niziurski, W., Meng, Y., Roux, B., and Berneche, S. Self-Learning Adaptive Umbrella Sampling Method for the Determination of Free Energy Landscapes in Multiple Dimensions. *Journal of chemical theory and computation*, 9(4):1885–1895, April 2013.

[41] Park, S., Kim, T., and Im, W. Transmembrane Helix Assembly by Window Exchange Umbrella Sampling. *Physical review letters*, 108(10):108102, March 2012.

[42] Efron, B. and Tibshirani, R. J. *An Introduction to the Bootstrap*. CRC Press, May 1994.

[43] Madden, D. R., Armstrong, N., Svergun, D., Perez, J., and Vachette, P. Solution X-ray Scattering Evidence for Agonist- and Antagonist-induced Modulation of Cleft Closure in a Glutamate Receptor Ligand-binding Domain. *Journal of Biological Chemistry*, 280(25):23637–23642, June 2005.

[44] Ahmed, A. H., Loh, A. P., Jane, D. E., and Oswald, R. E. Dynamics of the S1S2 Glutamate Binding Domain of GluR2 Measured Using 19F NMR Spectroscopy. *Journal of Biological Chemistry*, 282(17):12773–12784, April 2007.

[45] Landes, C. F., Rambhadran, A., Taylor, J. N., Salatan, F., and Jayaraman, V. Structural landscape of the isolated ligand binding domain of single AMPA receptors. *Nature chemical biology*, 7(3):168–173, March 2011.

Chapter 3 - High conformational variability in the GluK2 kainate receptor ligand-binding domain

Reproduced in part from a manuscript accepted at:

Structure

by

Tyler J. Wied, Alfred C. Chin, Albert Y. Lau

The kainate family of ionotropic glutamate receptors (iGluRs) mediates neurotransmission in pre- and post-synaptic neurons. Previously computed conformational potentials of mean force (PMFs) for iGluR bilobed ligand-binding domains (LBDs) revealed subtype-dependent conformational differences between AMPA and NMDA iGluR families. Here we report PMFs for GluK2, a member of the kainate receptor family, in apo and glutamate-bound states. Both apo and glutamate-bound GluK2 LBDs preferentially access a closed-cleft conformation. Apo GluK2 exhibits a surprisingly high degree of conformational flexibility, accessing both open and closed states. Comparing across iGluR subtypes, these results are most similar to the glycine-binding GluN1 and GluN3A subunits of the NMDA receptor family and differ from glutamate-binding GluA2 and GluN2A subunits. To test the thermodynamic contribution of cross-lobe interactions on the stability of closed-cleft LBDs, we computed PMFs for two GluK2 mutants, D462A and D656S. D462A, but not D656S, significantly weakens the closed-cleft conformation of the glutamate-bound LBD. Small-angle X-ray scattering (SAXS) experiments were carried out using the glutamate-bound GluK2 LBD and compared with a theoretical SAXS profile derived from the PMF. Weighting the theoretical SAXS calculation by the PMF, which represents a Boltzmann distribution of LBD conformations, improves agreement with the experimental result compared with a calculation using only the crystal structure of the LBD, thereby validating our free energy calculations. Principal component analysis (PCA) sheds light on the different conformational dynamics experienced by glutamate-bound and apo GluK2 during simulation, and these different modes of motion could be important for channel gating. Finally, an initial structural characterization of apo GluK2 is presented.

3.1 Background

Ionotropic glutamate receptors (iGluRs) mediate neurotransmission at excitatory synapses in the mammalian central nervous system [1]. Proper iGluR function is required for higher-order processes such as learning, memory, and synaptic plasticity, and their central role in cognition makes them attractive pharmaceutical targets [2]. The four primary iGluR families are the kainate receptors, α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptors, N-methyl-D-aspartic acid (NMDA) receptors, and δ receptors [3, 4]. Kainate receptors are expressed both pre- and post-synaptically, where they modulate cell depolarization, neurotransmitter release, and neuronal excitability [5, 6].

Functional iGluRs are tetrameric assemblies in which each protein subunit consists of an amino-terminal domain (ATD), a clamshell-like ligand-binding domain (LBD), a transmembrane pore domain (TMD), and a regulatory C-terminal tail (CTD) [7]. Receptor activation is associated with a series of coupled conformational changes that begin with the binding of an agonist to the LBD, which triggers LBD closure and concomitant opening of the ion channel pore. iGluRs exhibit a wide array of gating kinetics [4]. AMPA and kainate receptors generally exhibit rapid activation, deactivation and desensitization kinetics. NMDA receptors are much slower than AMPA and kainate receptors, activating on the order of tens to hundreds of milliseconds, and deactivating on timescales between tens and thousands of milliseconds, with weak desensitization. In addition to ion channel activation, glutamate binding and LBD dynamics mediate iGluR biogenesis. Conformational transitions between open and closed states of kainate receptors regulate plasma membrane expression and oligomeric assembly [8].

The conformational free energy landscapes, or potentials of mean force (PMFs), governing LBD closure for apo and natural agonist-bound proteins have been reported for AMPA and NMDA receptors but not for kainate receptors [9–12], although PMFs have been reported for a kainate receptor LBD bound to photoswitchable ligands [13, 14]. AMPA and NMDA receptor LBDs are characterized by a spectrum of qualitatively different PMFs. For instance, closed LBD conformations are energetically unfavorable for apo GluA2, preferring instead to sample open conformations, whereas apo GluN3A and GluN1 readily access both open and closed conformations. Glutamate-bound NMDA and AMPA receptors possess global free energy minima in a closed conformation, with energetically unfavorable open states, but the steepness of the PMFs between open and closed conformations varies among subtypes. In this study, we sought to determine whether computed PMFs for the LBD of a prototypical kainate receptor, GluK2, more closely resemble AMPA or NMDA receptors, or whether they are unique. This knowledge would advance our understanding of the structural thermodynamics underpinning iGluR function.

Here we report PMFs for apo and glutamate-bound GluK2 LBDs, computed using umbrella sampling molecular dynamics (MD) simulations. There are notable differences observed between the GluK2 PMFs and the PMFs of glutamate-binding NMDA and AMPA receptor subunits. Apo GluK2 is stable in both open and closed cleft conformations, unlike apo GluA2 and GluN2A, which preferentially access open conformations. One plausible factor contributing to the observed differences are cross-lobe interactions across the LBD cleft. Kainate receptor LBDs have several cross-lobe interactions that are not present in AMPA receptors, and these interactions have been reported to regulate the affinity for glutamate [15]. Additionally, GluK2 with cleft residues mutated to analogous GluA2 residues behave similarly to GluA2 in electrophysiology experiments, and vice versa with GluA2 residues mutated to analogous GluK2 residues [15]. To test the thermodynamic contribution of these cross-lobe interactions, we mutated two residues that participate in two sites of cross-lobe interaction, D656S and D462A, and computed the PMFs in the glutamate-bound and apo states. Our results predict a reduction in the stability of the D462A LBD in closed conformations, consistent with the functional assays. We compare experimental small-angle X-ray scattering (SAXS) results using the GluK2 LBD with a theoretical SAXS profile generated from a PMF-weighted ensemble of LBD-glutamate complex conformers. Next, we perform principal component analysis and show that apo and glutamate-bound GluK2 sample phase space differently. Whereas apo GluK2 experiences most of its motion hinge bending, glutamate-bound GluK2 experiences a variety of conformational motions. PCA might identify modes of motion important for channel gating beyond hinge bending. For instance, a twisting motion in GluA2 was previously shown to contribute significantly to GluA2 activation [16]. Finally, we present structural characterization of apo GluK2. We validate the absence of glutamate in solution by performing a protease protection assay as well as circular dichroism which shows slight changes to secondary structure. We then present initial attempts to characterize apo GluK2 by SAXS and X-ray crystallography.

3.2 Results

3.2.1 Conformational dynamics of the GluK2 LBD

To probe the conformational dynamics of the GluK2 LBD, we computed conformational PMFs for glutamate-bound and apo LBDs. The PMFs were computed using an umbrella sampling approach along a two-dimensional (2D) order parameter, (ξ_1, ξ_2), that reports the extent of openness of the LBD clamshell (**Fig. 3-1**). ξ_1 and ξ_2 each corresponds to a distance between the top lobe (Lobe 1) and the bottom lobe (Lobe 2) of the LBD. This approach has been previously employed to study large-scale conformational transitions in iGluR LBDs [9–12,17]. The glutamate-bound and apo LBDs share a global free energy minimum in the closed state at (ξ_1, ξ_2) = (10.1, 7.6 Å) and (10.5, 6.1 Å), respectively (**Fig. 3-2**). However, whereas the glutamate-bound LBD possesses only one free energy minimum, the apo LBD exhibits a broad free energy basin at more open conformations extending from (ξ_1, ξ_2) = (12, 10 Å) through (19, 20 Å). As expected, the LBD is freer to explore open conformations in the absence of a bound glutamate ligand (quantified below). This freedom, however, is surprisingly large, as the open-state basin remains below 2 kcal/mol, for the most part.

To test the contribution of cross-lobe interactions to the thermodynamic stability of the closed-cleft conformation, we disrupted these interactions by separately generating the cleft residue mutations D656S and D462A, where the amino acid substitutions are the analogous residues in GluA2. The PMF of the glutamate-bound D656S LBD resembles the PMF of its wild-type (WT) counterpart and features a global free energy minimum at (ξ_1, ξ_2) = (10.0, 7.2 Å), with increasing free energies at more open conformations (**Fig. 3-3**A). The PMF of the apo D656S LBD also looks broadly similar to the PMF of the WT apo LBD, with a global minimum at (ξ_1, ξ_2) = (10.0, 6.5 Å) (**Fig. 3-3**B). In the D656S PMFs, however, the free energy well in the immediate vicinity of the crystallographic closed state (see **Fig. 3-2**A) is narrower compared with the WT PMFs. This suggests the D656S mutation may slightly destabilize the crystallographic closed conformations. Overall, however, the thermodynamic effect of D656S appears to be small. In contrast, the D462A mutant significantly destabilizes closed-cleft conformations in the glutamate-bound LBD (**Fig. 3-3**C), with the free energy of open state configurations reaching only 1–2 kcal/mol compared with 4–5 kcal/mol for WT and D656S (**Fig. 3-2**A and **Fig. 3-3**A,C). Nonetheless, both apo and glutamate-bound D462A LBDs share a global free energy minimum near the crystallographic closed state [apo: (ξ_1, ξ_2) = (10.3, 6.1 Å), glutamate bound: (ξ_1, ξ_2) = (9.9, 6.0 Å)]. Results for apo D462A broadly resemble WT GluK2, suggesting minimal effect of the mutation.

On the other hand, the open states of the glutamate-bound D462A LBD are substantially more favorable than either the glutamate-bound WT or D656S LBDs (**Fig. 3-3**C). This suggests D462 plays an important role in stabilizing closed-cleft conformations. The PMF of the apo D462A LBD is similar to that of the apo WT LBD (**Fig. 3-3**D). Presumably, since the apo LBD is already highly flexible, the D462A mutation produces little effect in the apo LBD. Error analysis for all PMFs are provided in **Fig. 3-4**.

How do the WT GluK2 LBD PMFs compare with the PMFs previously computed for AMPA and NMDA receptor LBDs in [9, 11]? Superposition of 1D projections of the 2D PMFs into an order parameter called ξ_{12} (see **section 3.4**) is a convenient way to compare PMFs (**Fig. 3-5**). The iGluR subtypes are structurally similar but not identical, which permits the comparison in ξ_{12}, albeit imperfectly. The PMF of the glutamate-bound GluK2 LBD resembles the PMFs for agonist-bound GluA2, GluN1, GluN2A, and GluN3A LBDs in that the global free energy minimum resides in the closed state, and the free energy increases in more open conformations (**Fig. 3-5**A). The slopes of the PMFs vary by iGluR subtype, with glutamate-bound GluA2 being the steepest and GluK2 being the shallowest. In other words, GluK2 exhibits the lowest free energy barriers to accessing open conformations. The PMF of the glycine-bound GluN1 closely resembles that of GluA2, whereas the slopes of both the glutamate-bound GluN2A and the glycine-bound GluN3A are in between that of GluA2 and GluK2. The bi-stable PMF of the apo GluK2 LBD is most similar to the glycine-binding NMDA receptor subtypes GluN1 and GluN3A (**Fig. 3-5**B). The PMFs of apo GluA2 and GluN2A are different in that they contain one broad free energy basin in open conformations and are relatively unstable in closed conformations.

3.2.2 Free energy difference between open and closed GluK2 LBDs

How can one define whether a given GluK2 LBD conformation is open or closed? One approach to differentiate these states is to define closed states as those that do not permit entry or exit of a glutamate ligand into or out of its binding pocket. The program CAVER [18] was used to search for possible passageways for glutamate between the protein surface and the binding cavity, given glutamate's estimated van der Waals radius of 3.22 Å [19]. Multiple LBD conformers generated by the simulations were evaluated, and a single access tunnel was identified. Y457 in Lobe 1 and V654 in Lobe 2 were identified as "bottleneck" residues that form the tightest constriction point along the tunnel (**Fig. 3-6**A). In a closed state, these bottleneck residues prevent the passage of glutamate into or out of the LBD. These residues are in nearly analogous locations as bottleneck residues previously identified in a ctenophore iGluR LBD [17]. At least 12 Å between the Y457 and V654 C_γ atoms is estimated to be required for glutamate to access its binding site. Therefore, conformations, i.e., (ξ_1, ξ_2), with bottleneck distances greater than 12 Å are classified as open, and all others are classified as closed. Plots of average bottleneck distance vs. (ξ_1, ξ_2) for WT, D656S, and D462A LBDs are shown in **Fig. 3-6**B-G. The values in each plot were calculated using ∼120,000 snapshots extracted from the simulations.

In an equilibrium ensemble of LBD conformations, what fraction is open vs. closed? The relative probability of the LBD occupying each state is determined by the ratio of the partition functions for each state,

$$\frac{\Omega_{\text{Open}}}{\Omega_{\text{Closed}}} = \frac{\int_{\text{Open}} e^{-W(\xi_1,\xi_2)/k_BT} d\xi_1 d\xi_2}{\int_{\text{Closed}} e^{-W(\xi_1,\xi_2)/k_BT} d\xi_1 d\xi_2} \tag{3.1}$$

where $W(\xi_1, \xi_2)$ is the PMF, and the region of integration corresponds to either (ξ_1, ξ_2) in which the bottleneck distance $\geq$12 Å (open) or $<$12 Å (closed); k_B is Boltzmann's constant, and T is temperature. The conformational free energy difference between open and closed states is then given by,

$$\Delta G_{\text{conf}} = -k_BT \ln \left(\frac{\Omega_{\text{open}}}{\Omega_{\text{Closed}}} \right) \tag{3.2}$$

The effect of binding a glutamate ligand on the open vs. closed LBD population is given by,

$$\Delta\Delta G_{\text{conf}} = \Delta G^{\text{conf}}_{\text{glu}} - \Delta G^{\text{conf}}_{\text{apo}} \tag{3.3}$$

Values of ΔG_{conf} and $\Delta\Delta G_{\text{conf}}$ are provided in Table 1. Bound glutamate stabilizes the closed cleft WT GluK2 LBD by 2.8 kcal/mol relative to its apo state, 2.6 kcal/mol for D656S, and 0.4 kcal/mol for D462A. The D462A mutation almost eliminates the difference between the glutamate-bound and apo states, suggesting that D462 plays an important role in stabilizing the closed cleft when glutamate is bound. In contrast, the D656S mutation has only a marginal effect on the stability of glutamate-bound and apo LBDs.

3.2.3 Small-angle X-ray scattering

How accurate are our computed conformational PMFs? To address this question, we compared theoretical small-angle X-ray scattering (SAXS) profiles generated using our PMFs with experimental SAXS data. A similar comparison has previously been carried out involving the GluA2 LBD in complex with several ligands [9]. Theoretical SAXS profiles were calculated, using CRYSOL [20], for 1369 simulated configurational snapshots spanning (ξ_1, ξ_2). A single Boltzmann-weighted average profile was calculated from each of the profiles, where the weights are given by the computed

PMF (see **section 3.4**). Additionally, the theoretical SAXS curve was calculated for the crystal structure of the glutamate-bound GluK2 LBD (PDB: 1S50) [21]. These theoretical profiles can be directly compared with experimental SAXS data, which represent an ensemble protein conformational distribution in solution [22]. **Fig. 3-7** shows a comparison of the theoretical SAXS profile, calculated using the glutamate-bound LBD PMF, with experimental SAXS data, collected using the GluK2 LBD in the presence of 10 mM glutamate, a concentration at which ∼100% of LBDs should be bound to glutamate [21].

The theoretical and experimental SAXS curves were scaled to I_0 and compared by computing the chi-squared statistic, χ^2. There is good agreement between the experimental and theoretical SAXS curves from simulation, $\chi^2 = 3.0$, which is a better overall fit to the experimental data than the theoretical SAXS curve from the crystal structure alone, $\chi^2 = 17.8$. There is also good agreement between the radii of gyration, R_g, with experimental $R_g = 19.9$ Å and theoretical $R_g = 19.6$ Å, which are both larger than the crystal structure $R_g = 19.1$ Å. These results suggest that our simulations accurately capture the populations of relevant LBD conformations that are not reflected in the crystal structure alone.

3.2.4 Structural characterization of apo GluK2

Apo GluK2 was purified by exhaustive dialysis (∼10^{20}-fold serial dilution) into solutions not containing L-glutamate. Apo LBDs can be distinguished from glutamate-bound GluK2 via protease protection assay [23] and circular dichroism (**Fig. 3-8**). Proteolysis protection assays were carried out for GluK2 in the presence of the trypsin protease. Glutamate-bound GluK2 is resistant to trypsin digestion after 120 minutes (**Fig. 3-8**A, *top* and *bottom*). Apo GluK2 LBD, in contrast, is vulnerable to trypsin digestion, and is nearly completely digested after 120 minutes (**Fig. 3-8**A, *middle*). Additionally, small differences are observed between apo and glutamate-bound GluK2 in their circular dichroism (CD) spectra, which is used to determine protein secondary structure character. Relative to apo GluK2, glutamate-bound GluK2 has more β-I character. This suggests the LBD loses some of its beta sheet character in its apo form, perhaps in the hinge, which undergoes the most structural change between open and closed conformations. We attempted to carry out SAXS experiments using the apo GluK2 LBD, but the protein persistently aggregated during our measurements.

There is no crystal structure of an apo GluK2 LBD. Preliminary attempts to grow apo GluK2 crystals generated promising hits, but we were ultimately unsuccessful in producing crystals viable for structure determination. Promising crystal hits are shown in **Fig. 3-9**. A complete summary of crystallization attempts are detailed in **section 3.4**. Generally, the vast majority of drops containing 7.5 mg/mL protein formed precipitates. Notably, conditions that produce crystals for glutamate-bound GluK2 [21] do not produce crystals for apo GluK2.

3.2.5 Principal component analysis

To assess large-scale conformational motions in our simulations, principal component analysis (PCA) was performed on apo and glutamate-bound GluK2 LBDs **Fig. 3-10**. For both apo and glutamate-bound GluK2, PC1 is a hinge bending, which corresponds to the canonical opening/closing of the LBD. PC2 for the apo LBD is a 'sweeping' motion, whereas for glutamate-bound LBD it is a 'rocking' motion. PC3 for the apo form is the 'rocking' motion and for the glutamate-bound form PC3 corresponds to the 'sweeping' motion. Notably, 87% of the variance is contained within PC1 alone for apo GluK2, and 93.2% of the variance is contained within PC1-3. The glutamate-bound LBD contains 28.1% of its variance within PC1, and only 56% of its variance within PC1-3. This suggests that motions of the apo form are dominated by the hinge-bending motion, whereas this motion is attenuated heavily in the glutamate-bound LBD. It is possible that the non-hinge bending motions observed here, such as the "sweeping" or "rocking" motions observed in PC1-2, play a role in the conformational dynamics of full-length receptor assemblies.

3.3 Discussion

To assess the conformational thermodynamics of GluK2 LBDs, we computed PMFs for GluK2 LBDs in apo and holo states. The glutamate-bound GluK2 LBD favors a closed-cleft conformation, whereas apo GluK2 is highly flexible and can access a broad region of conformational space spanning closed and open conformations. The PMF of glutamate-bound GluK2 resembles the PMFs of other agonist-bound iGluR LBDs, but with a shallower free energy profile in open conformations. Apo GluK2 is unique in that it can access both closed and open conformations – other apo iGluRs that bind glutamate do not access closed conformations with a high probability. To quantify the free energy difference between open and closed LBD states, we classified conformational snapshots into "open" and "closed" states based on a bottleneck analysis of ligand-binding tunnels. This approach allows us to calculate the effect glutamate has on the relative stability of open and closed LBD conformations. It should be noted that the free energy calculations presented here constitute only a partial contribution to the total free energy of glutamate binding, as ligand docking is not taken into account. For a complete treatment of ligand-binding free energies to an iGluR LBD, see [10].

Why can apo GluK2 easily access a closed cleft conformation, but apo GluA2 cannot? One possibility is the presence of cross-lobe interactions not present in the GluA2 LBD. To test the thermodynamic effect of disrupting inter-cleft interactions, we computed PMFs for two mutants, D656S and D462A, in which the substituted amino acids are the analogous residues in GluA2. We found that D462A, but not D656S, significantly destabilizes the closed-cleft LBD, effectively stabilizing the open LBD. Notably, it was the PMF of the glutamate-bound, and not apo, D462A LBD that was most significantly affected. Apo D462A GluK2 still visits the closed state with a high relative probability.

Another explanation for the LBD's thermodynamic behavior is bulk electrostatic attraction or repulsion between the top and bottom lobes, as has been discussed previously for GluA2 [24]. The GluA2 LBD has a total net charge of $+5$, where the top lobe of has a $+4$ charge, and the bottom lobe has a $+1$ charge, hence the two lobes should experience net repulsion. The GluK2 LBD upper lobe has a net charge of -2, and the bottom lobe has a net charge of 0, suggesting GluK2 experiences less cross-lobe repulsion than GluA2. More specifically, repulsion between protein surfaces of the upper and lower lobes might predict the preference for apo LBDs to visit closed cleft conformations. **Fig. 3-11** and **Fig. 3-12** show electrostatic surfaces for various iGluR LBDs both from the "front" of the LBD and the surfaces of the upper and lowers that face the binding cleft. From these views, deep blue pockets can be observed, corresponding to the glutamate/glycine binding pocket. However, other than that, the surfaces are heterogeneous, and it is difficult to draw general conclusions from these surfaces. A Poisson-Boltzmann calculation or perhaps the CHARMM Block module, which allows users to tune what forces groups of atoms "feel", would reveal the Coloumbic contribution to the PMFs presented here. We attempted to use the CHARMM Block module to address this question, and encountered strange behavior. Specifically, with our configuration of Block, it causes an acceleration of the protein in the Z dimension. This does not appear to be dependent on choice of integrator or ensemble (NPT vs NVT).

To compare simulation results with experiment, we performed SAXS on the GluK2 LBD and computed theoretical SAXS curves from our simulated ensemble. Overall, the agreement between the simulation and experiment are an improvement over a comparison with the crystal structure alone, suggesting the glutamate-bound GluK2 LBD accesses an array of conformational states not represented by the crystal structure. Previous comparisons of LBD PMFs with SAXS [9] and single-molecule FRET (smFRET) [25] studies have also shown good agreement between experimental and computational results.

Here the intrinsic conformational thermodynamics of a monomeric GluK2 LBD is studied. Future studies will address the conformational dynamics of LBD dimer assemblies as well as intact, tetrameric receptors.

3.4 Methods

3.4.1 Molecular dynamics simulations

Molecular dynamics simulations were carried out in CHARMM [26] and NAMD [27] using the TIP3P water model [28]. The GluK2 LBD was chosen as a representative kainate receptor LBD (PDB: 1S50, chain A) [21]. The N-terminal glycine residue not modelled in the crystal structure was built using MODELLER [29]. Crystallographic waters buried within the LBD were included in the simulations, whereas waters outside the LBD were removed. A 90 × 70 × 60 Å^3 orthorhombic water box with ~150 mM NaCl was used for all simulations; the number of ions was adjusted to set the net charge of the system to zero (apo: $\text{Na}^+ = 30$, $\text{Cl}^- = 28$; glutamate-bound: $\text{Na}^+ = 30$, $\text{Cl}^- = 27$). In simulations of glutamate-bound LBD, the ligand was prevented from dissociating from the binding pocket by using a half-harmonic restraint to keep the α-carboxyl oxygens of the ligand within 2.8 Å of the guanidine group of R492. Initial conformations for umbrella sampling were generated using biasing potentials [9]. Equilibration took place in the NVT ensemble using applied restraints on backbone and sidechain atoms, which were slowly released during equilibration. Production runs took place in the NPT ensemble at 1 atm and 300 K [30]. Electrostatic interactions were computed using the particle mesh Ewald (PME) algorithm [31]. Umbrella sampling was performed using a 2D order parameter, (ξ_1, ξ_2). The first dimension, ξ_1, is defined as the distance between the center-of-mass of all non-hydrogen atoms in residues 90–92 in the top lobe and the center-of-mass of non-hydrogen atoms in residues 142–143 in the bottom lobe. The second dimension, ξ_2, is defined as the distance between the center-of-mass of non-hydrogen atoms in residues 12–14 in the top lobe and the center-of-mass of non-hydrogen atoms in residues 174–175 in the bottom lobe. A force constant of 2 kcal/mol/Å^2 was used for ξ_1 and ξ_2. This order parameter spans the relevant regions of conformational space surrounding the closed- and open-cleft conformations observed in crystal structures.

3.4.2 Free energy calculations

126 windows were used for umbrella sampling, and each window was sampled for 1.8 ns, for a total of 226.8 ns. Umbrella sampling distribution functions were unbiased and recombined using the weighted histogram analysis method (WHAM) [32,33].

1D projections of the 2D PMFs were generated by computing the PMF along a hybrid order parameter that is the average of ξ_1 and ξ_2, which we call ξ_{12}. This is done by mapping each ξ_1 and ξ_2 pair to a bin in ξ_{12}, with a bin width of 0.1 Å, and then summing the values from the 2D unbiased distribution in each bin to create a 1D unbiased distribution. The 1D PMF, W(ξ_{12}), is given by,

$$\mathrm{W}(\xi_{12}) = -k_B T \ln \left(\frac{\rho(\xi_{12})}{\sum_{j=1}^{N} \rho_j(\xi_{12})} \right) \tag{3.4}$$

where is the 1D unbiased distribution function, and is the partition function.

Uncertainty in the simulation results was assessed by block averaging [34]. Briefly, the simulation trajectory for each umbrella sampling window was split into ten equally sized "blocks", and WHAM was performed on each of the blocks to generate ten block PMFs. The standard deviation of the block PMFs was then computed.

3.4.3 Expression and purification of the GluK2 LBD

The GluK2 LBD construct was provided by Mark Mayer (NIH) on a pET-22b plasmid (Novagen), with the same sequence as previously described [21]. The GluK2 LBD expresses as a soluble, monodisperse protein in BL21(DE3) E. coli (Novagen). Expression was carried out in cultures of Terrific Broth grown to OD600 = 0.7 in the presence of 100 μ g/mL Ampicillin and induced with 400 μM IPTG overnight at 4°C. Purification was performed on a Ni^{2+} NTA column (Roche), followed by thrombin cleavage and size exclusion chromatography. For SAXS and the proteolysis protection assay, purified protein was dialyzed into a solution containing 20 mM NaCl, 2 mM Tris, pH 8, 10 mM L-glutamate, 1% glycerol, and 1 mM TCEP.

3.4.4 Small-angle X-ray scattering

SAXS experiments were performed using an in-house Rigaku BioSAXS-2000 system. Different concentrations of GluK2 LBD (0.5, 1, 2, 5 mg/mL) were suspended in a buffer containing 20 mM NaCl, 2 mM Tris, pH 8, 10 mM L-glutamate, 1% glycerol and 1 mM TCEP. Samples were loaded into the capillary by an automatic sample changer. Buffer subtraction and data merging were performed using SAXSLab (Rigaku).

Theoretical SAXS curves were generated using CRYSOL [35],with the solvent density value set to 0.36 electrons/Å^3, which lies within an acceptable range, as previously described [20]. A total of 1369 randomly-selected snapshot conformations were used to generate an ensemble of SAXS curves. A Boltzmann weight determined by the corresponding PMF value for each conformation was applied to the output intensity values. These weighted SAXS curves were then combined into a final Boltzmann-weighted ensemble-average SAXS profile. The ensemble-average scattering profile is given by,

$$\langle I_{\text{eq}}(q)\rangle = \frac{\sum_j I_{\text{sol,j}}(q)e^{-W_j(\xi_1,\xi_2)/k_BT}}{\sum_j e^{-W_j(\xi_1,\xi_2)/k_BT}} \tag{3.5}$$

where $\mathrm{I}_{sol}(q)$ is the intensity curve, $\mathrm{W}_j(\xi_1, \xi_1)$ is the PMF, k_B is Boltzmann's constant, and T is temperature. The Guinier approximation was used in the range where $qR_g < 1.3$ to calculate the radius of gyration, R_g, and the forward scattering, I_0, from the following fit,

$$\ln(I_{\text{eq}}(q)) = \ln(I_0) - \left(\frac{R_g^2}{3}\right)q^2 \tag{3.6}$$

where the slope and y-intercept of a plot of [ln(I(q)) vs. q^2] give the R_g and I_0, respectively. The χ^2 function was used to assess the agreement between the theoretical and experimental scattering profiles,

$$X^2 = \frac{1}{N-1}\sum_{m=1}^{N}\left(\frac{I_{\text{exp}}(q_m) - I_{\text{calc}}(q_m)}{\sigma(q_m)}\right)^2 \tag{3.7}$$

where N is the number of data points in the scattering profile, $I_{\text{exp}}(q_m)$ and $I_{\text{calc}}(q_m)$ are the experimental and calculated scattering profiles, respectively, and $\sigma(s_i)$ is the experimental standard deviation. Interpolation of data was performed using a cubic spline.

3.4.5 Structural characterization of apo GluK2

Apo GluK2 is obtained by exhaustive dialysis ($\sim 10^{20}$-fold serial dilution) into a buffer containing 20 mM NaCl, 2 mM Tris, pH 8, and 1% glycerol. The dialyzed protein solution was run over a size-exclusion chromatography column to remove aggregated protein. Limited proteolysis protection assays were performed with 1 mg/mL purified apo and glutamate-bound GluK2 with the buffer composition described above, as well as purified apo GluK2 with 10 mM L-glutamate re-added just prior to the assay. Trypsin was added (1/20 wt/wt ratio) to room temperature protein samples, and reaction aliquots were collected at 30, 60, and 120 minutes for SDS-PAGE. Control reactions were performed in the absence of trypsin as well as excess trypsin. Circular dichroism (Aviv Model 400 CD spectrometer) was carried out between 190 and 300 nm in buffers containing 20 mM NaCl, 2 mM Tris, pH 8, 2% glycerol, 0.25 mM TCEP, and 10 mM L-glutamate for glutamate-bound GluK2.

Crystal screens for apo GluK2 LBD were carried out with the JCSG Core Suite 1 at various protein concentrations (5 and 7.5 mg/mL) and temperatures (4 °C and 20 °C). Apo GluK2 was suspended in a buffer containing 20 mM NaCl, 2 mM Tris, pH 8, and 1 mM TCEP. Screening was carried out with the hanging drop method with a mosquito LCP, which combined 300 nL of protein sample with 300 nL of reservoir solution.

3.4.6 Principal component analysis

Principal component analysis (PCA) was carried out with the Bio3D package for R ([36]) as previously described in [11]. In brief, a subset ensemble was generated from the C_α carbons of umbrella sampling conformations within 2 kcal/mol of the global PMF free energy minima. This subset ensemble was used to generated the covariance matrix from which the eigenvectors and eigenvalues of PCA are computed. We applied an approximate Boltzmann-weight to the selected conformations: conformations between 0-1 kcal/mol were duplicated five times relative to the conformations between 1-2 kcal/mol.

3.5 Acknowledgements

We thank Mark Mayer for providing the GluK2 LBD plasmid. We also thank Michael Love, Rick Hooy, and John Belcher for assistance with SAXS data collection and analysis. Computational resources were provided by the Maryland Advanced Research

Computing Center (MARCC) at Johns Hopkins University. This work was supported by National Science Foundation Graduate Fellowship 1232825 (to T.W.) and National Institutes of Health Grant R01GM094495 (to A.Y.L.).

3.6 Author Contributions

T.W. and A.Y.L. designed the project. T.W. and A.C.C. performed the research. T.W., A.C.C., and A.Y.L. analyzed the data and wrote the manuscript.

3.7 Declaration of interests

The authors declare no competing interests.

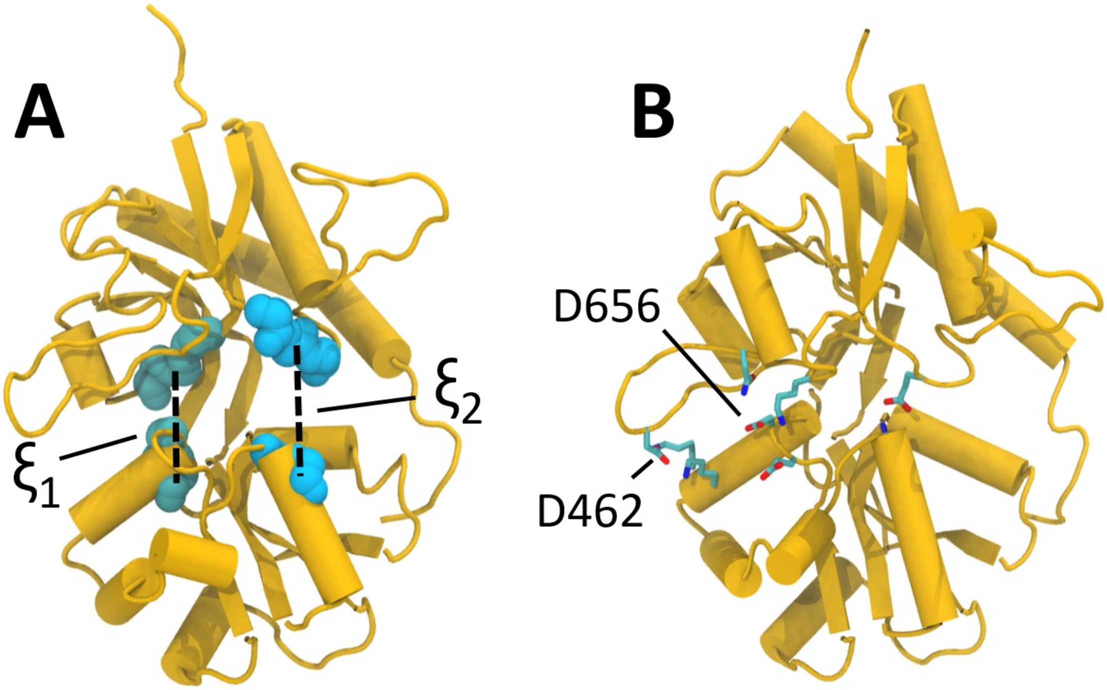


Fig. 3-1: **GluK2 ligand-binding domain structure.** (A) A 2D order parameter, (ξ_1, ξ_2), describes the conformational state of the LBD in our simulations. ξ_1 and ξ_2 each indicate a distance, represented by dashed lines, between the residues shown in spheres. (B) Residues that participate in interactions between the top and bottom lobes are depicted. Residues chosen for mutational analysis, D656 and D462, are labeled.

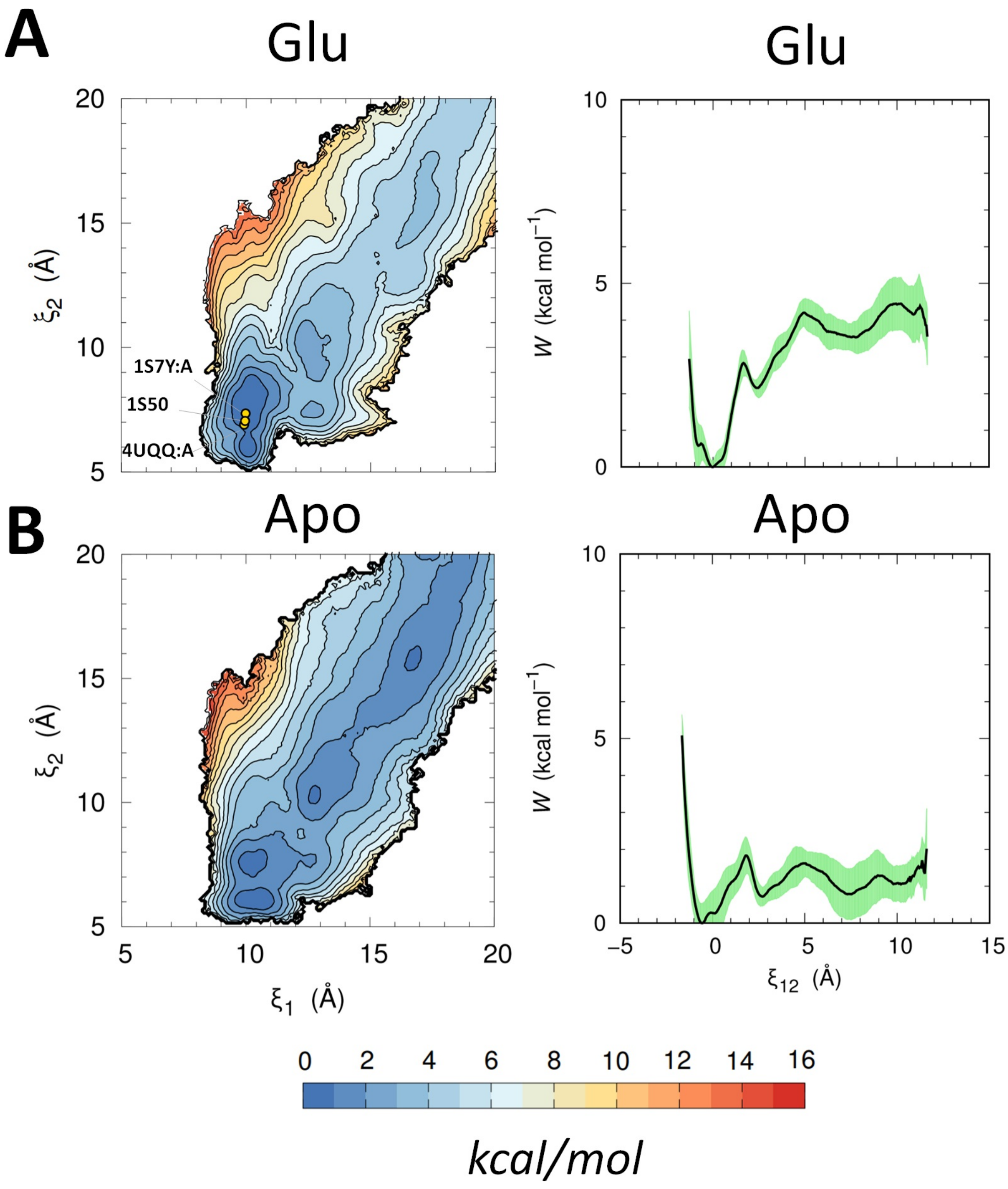


Fig. 3-2: **Potential of mean force (PMF) calculations for WT GluK2 LBDs.** (A) Glutamate-bound. (B) Apo. 2D PMFs are on the left, and 1D PMFs, which are projections of the 2D PMFs according to $\xi_{12} = (\xi_1 + \xi_2)/2$, are on the right. The 2D PMFs are contoured by 1 kcal/mol increments. ξ_{12} for the 1D PMFs are adjusted so the ξ_{12} value is set to 0 Å where the glutamate-bound LBD has its global free energy minimum. Standard deviation is shown in green and is calculated using block averaging.

Fig. 3-3: PMFs for mutant GluK2 LBDs. (A) D656S-glutamate, (B) apo D656S, (C) D462A-glutamate, and (D) apo D462A. 2D plots are on the left, and 1D projections are on the right. Standard deviation is shown in green and is calculated using block averaging.

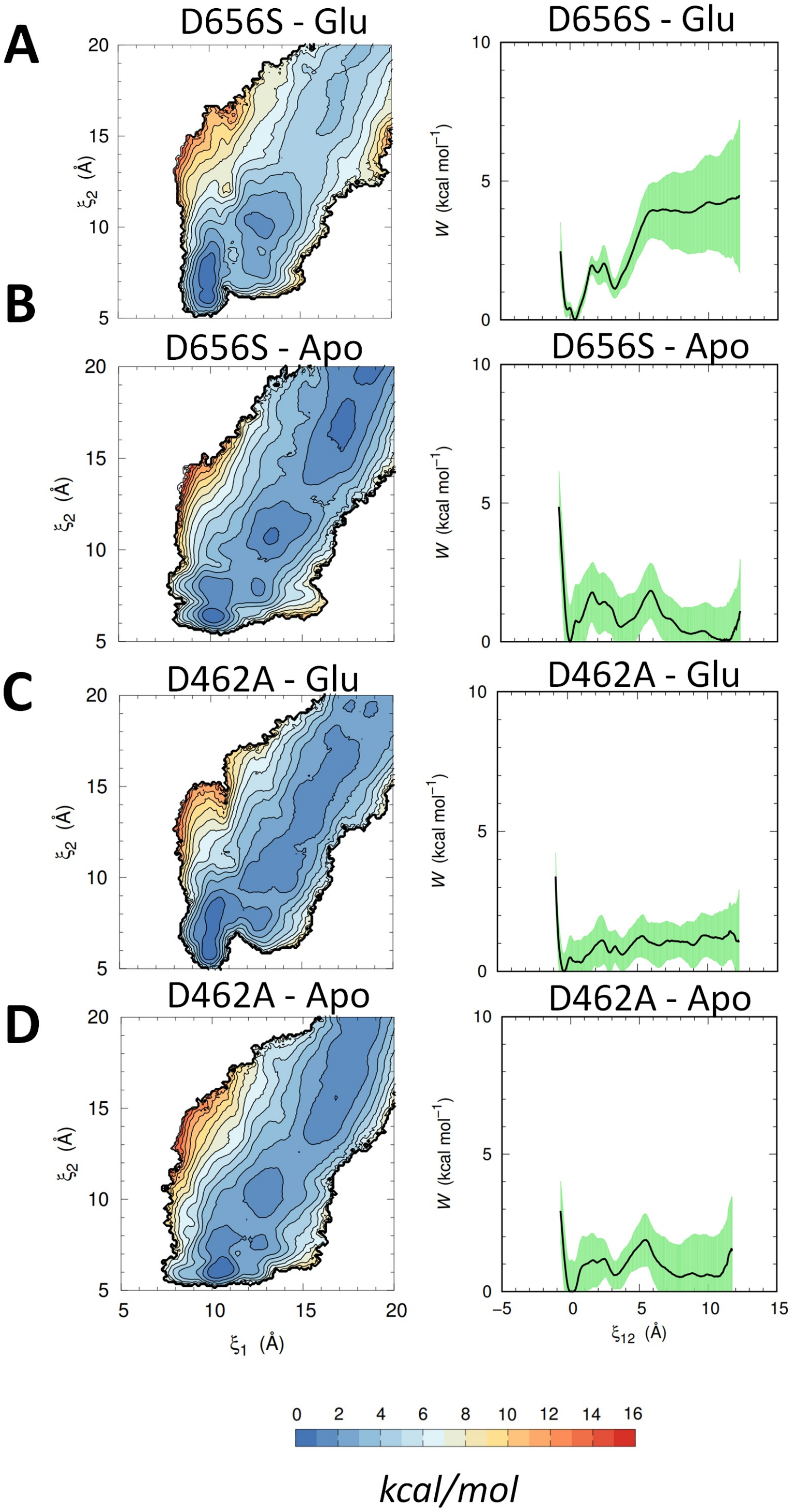

A
D656S - Glu
D656S - Glu
B
D656S - Apo
D656S - Apo
C
D462A - Glu
D462A - Glu
D
D462A - Apo
D462A - Apo
ξ2 (Å)
ξ1 (Å)
W (kcal mol−1)
ξ12 (Å)
5
10
15
20
−5
0
0
2
4
6
8
10
12
14
16
kcal/mol

Fig. 3-3

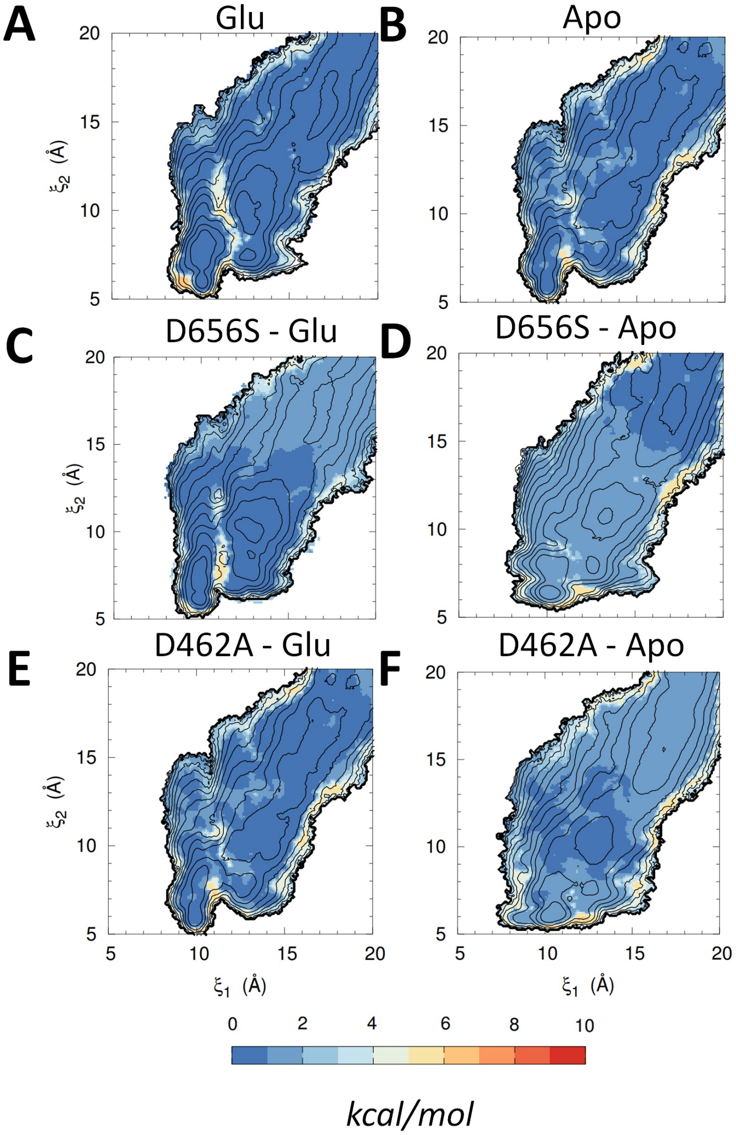


Fig. 3-4: **GluK2 LBD PMF standard deviations.** (A) WT-glutamate, (B) apo WT, (C) D656S-glutamate, (D) apo D656S, (E) D462A-glutamate, and (F) apo D462A. Contour lines correspond to the PMF, and colors correspond to the standard deviation, calculated using block averaging.

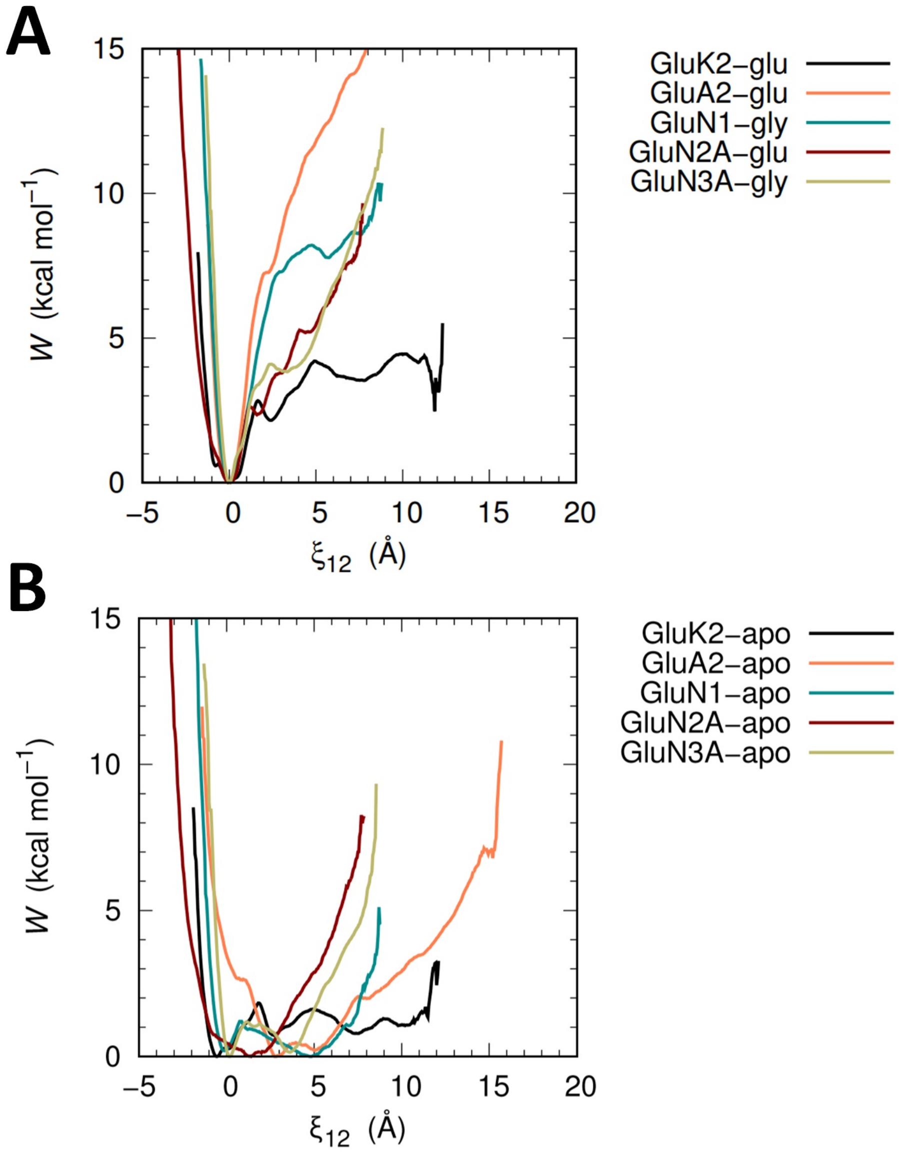


Fig. 3-5: **Comparison of 1D PMFs for AMPA, NMDA, and kainate receptors.** (A) Agonist-bound PMFs for GluK2-glutamate (black), GluA2-glutamate (orange), GluN1-glycine (turquoise), GluN2A-glutamate (red), and GluN3A-glycine (gold). (B) Apo PMFs for GluK2 (black), GluA2 (orange), GluN1 (turquoise), GluN2A (red) and GluN3A (gold). The PMFs of the various LBDs are aligned such that the order parameter, ξ_{12}, is set to a value of 0 Å where the agonist-bound LBD has its global free energy minimum.

Fig. 3-6: Bottleneck analysis. (A) Residues defining the bottleneck to the glutamate binding site, Y457 and V654. (B-G) The average bottleneck distance at each (ξ_1, ξ_2) is overlaid with the corresponding PMF contours. The color corresponds to the average bottleneck distance. An average bottleneck distance of greater than 12 Å is considered an open conformation. The dashed boxes in the upper right corners depict regions of conformational space that roughly, but not always, correspond to open LBD clefts.

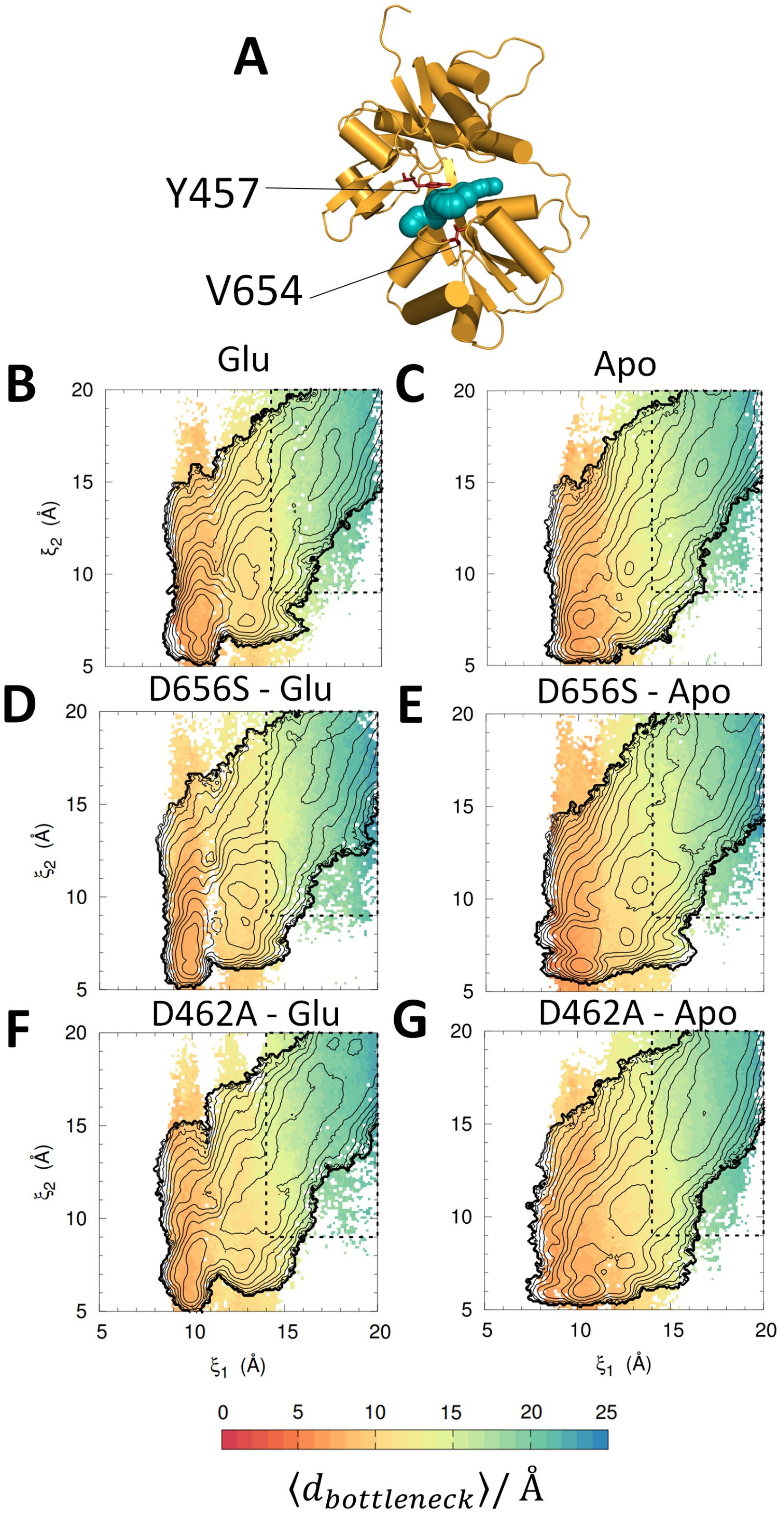

A
Y457
V654
B
Glu
C
Apo
D
D656S - Glu
E
D656S - Apo
F
D462A - Glu
G
D462A - Apo
20
15
10
5
5
10
15
20
ξ_2 (Å)
ξ_1 (Å)
0
5
10
15
20
25
$\langle d_{bottleneck} \rangle$/ Å

Fig. 3-6

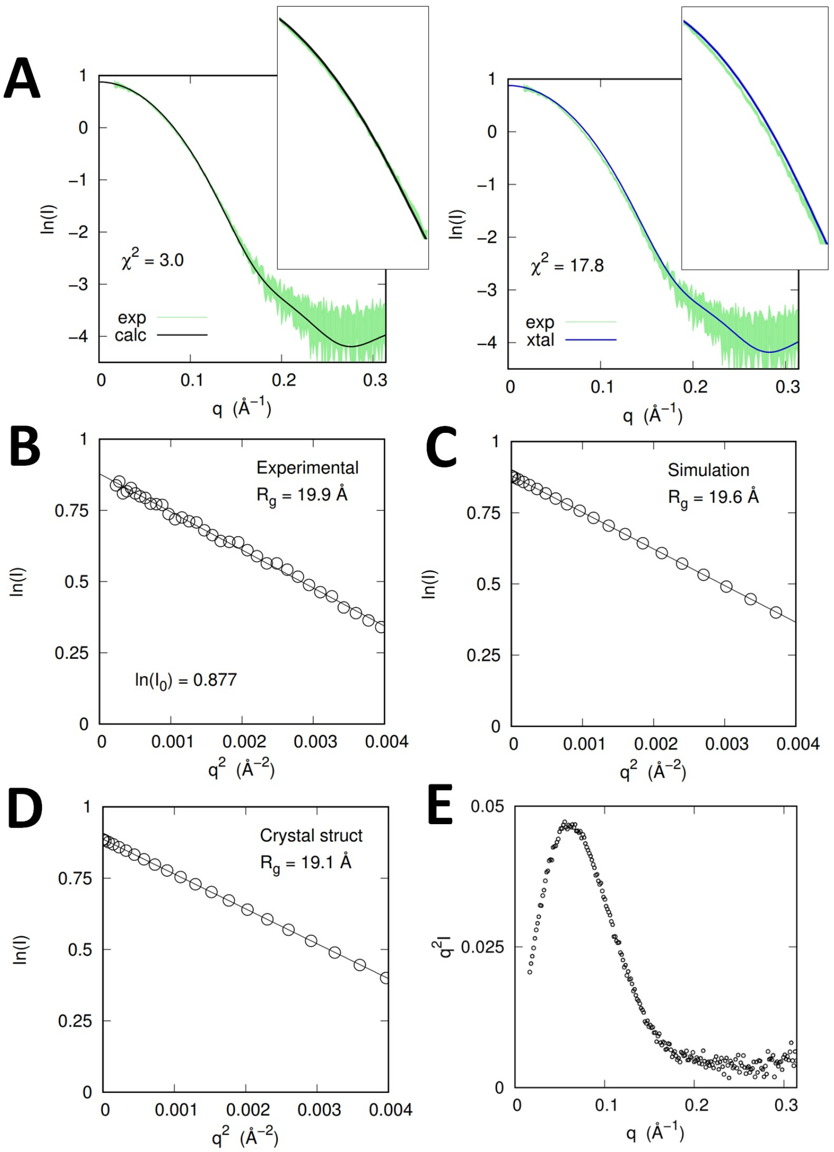


Fig. 3-7: **Comparison of experimental and theoretical small angle X-ray scattering (SAXS) of GluK2 bound to glutamate.** (A) SAXS log intensity plot with an overlay of experimental results (green), theoretical results from MD simulation (black, left), and theoretical results from the crystal structure alone (blue, right). The inset boxes in the upper right corners are zoom-ins of the intensity plot from roughly $0.05 < q < 0.15$ Å^{-1}. (B) Guinier plot for experimental results (R_g = 19.9 Å). (C) Guinier plot for theoretical results from MD simulation (R_g = 19.6 Å). (D) Guinier plot for theoretical results from the crystal structure alone (R_g = 19.1 Å). (E) Kratky plot with its characteristic shape indicating a folded protein.

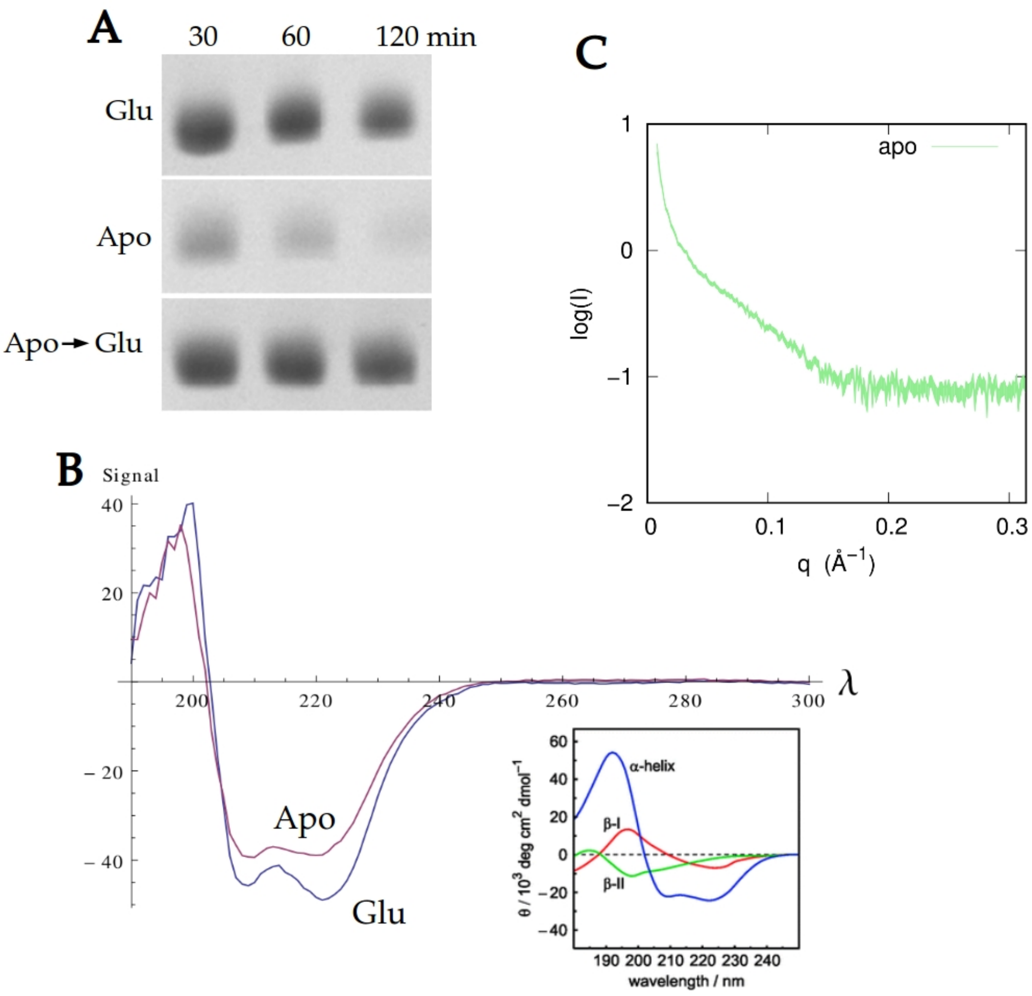
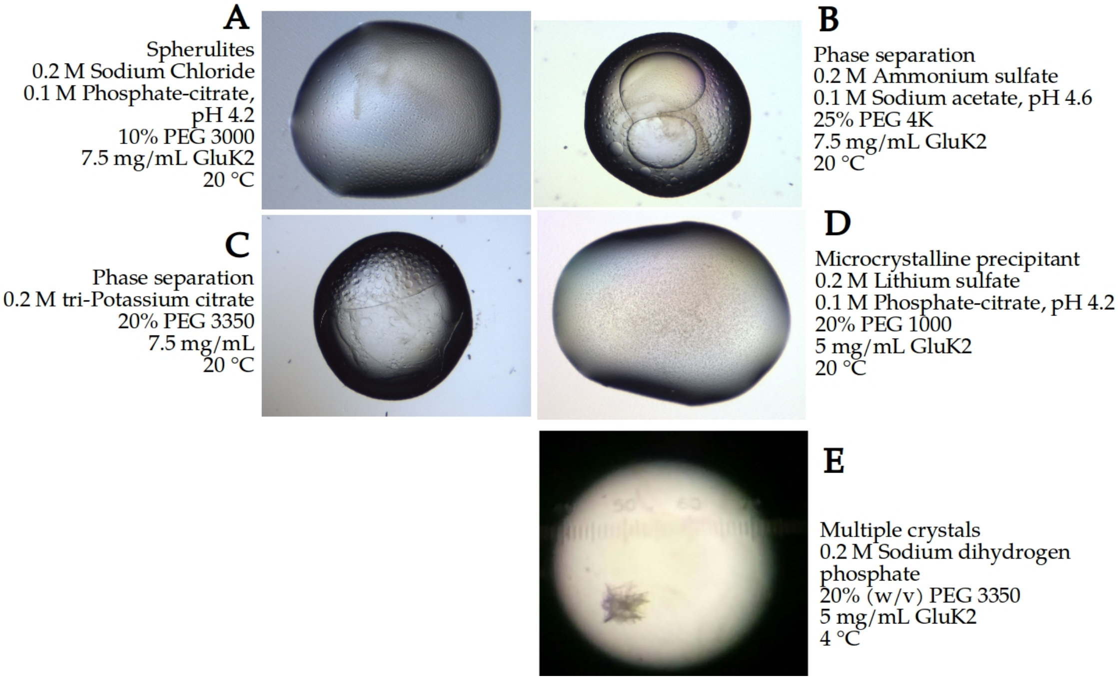


Fig. 3-8: **Validation of apo GluK2 and SAXS attempts.** (A) Proteolysis protection assay of GluK2 LBD in the presence of trypsin. Glutamate-bound (*top* and *bottom*) GluK2 is resistant to trypsin after 120 minutes. Apo GluK2 (*center*) is almost completely degraded after 120 minutes. (B) Circular dichroism suggests subtle secondary structure differences between apo and glutamate-bound GluK2. Glu-bound GluK2 appears to have more β-I character than apo GluK2. (C) Apo GluK2 SAXS results. These results are strongly indicative of aggregation, even at low (0.5 mg/mL) concentrations.



A
Spherulites
0.2 M Sodium Chloride
0.1 M Phosphate-citrate, pH 4.2
10% PEG 3000
7.5 mg/mL GluK2
20 °C

B
Phase separation
0.2 M Ammonium sulfate
0.1 M Sodium acetate, pH 4.6
25% PEG 4K
7.5 mg/mL GluK2
20 °C

C
Phase separation
0.2 M tri-Potassium citrate
20% PEG 3350
7.5 mg/mL
20 °C

D
Microcrystalline precipitant
0.2 M Lithium sulfate
0.1 M Phosphate-citrate, pH 4.2
20% PEG 1000
5 mg/mL GluK2
20 °C

E
Multiple crystals
0.2 M Sodium dihydrogen phosphate
20% (w/v) PEG 3350
5 mg/mL GluK2
4 °C

Fig. 3-9: **Preliminary crystal screen for apo GluK2 LBD.** Shown in *A*-*E* are the most promising hits from JCSG Core Suite 1 with 5-7.5 mg/mL protein at either 4 °C or 20 °C. GluK2 LBD. Spherulites are seen in *A*, phase separation in *B* and *C*, microcrystals/precipitant in *D*, and a single collection of crystal shards in *E*.

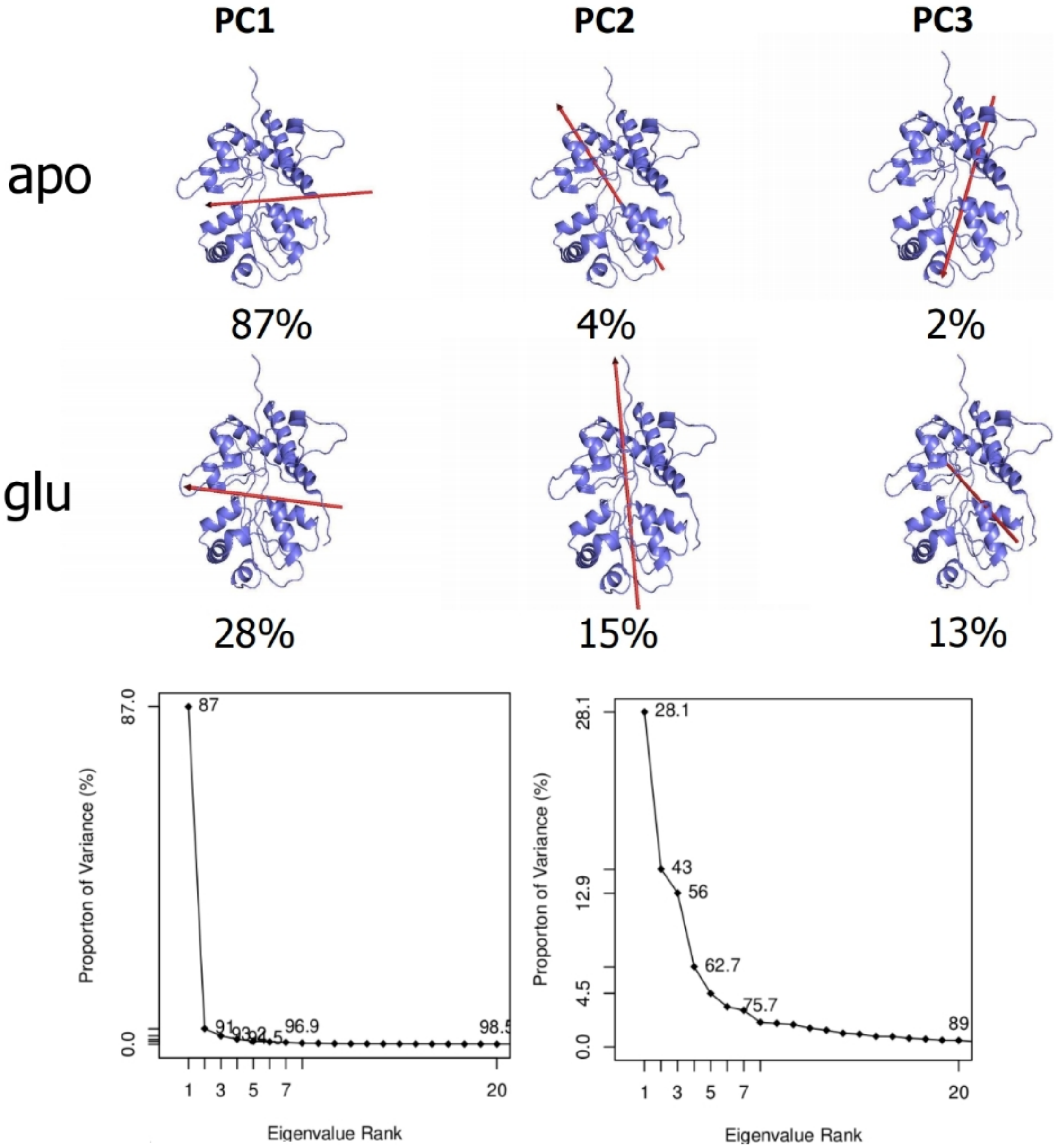


Fig. 3-10: **Principal component analysis of GluK2 LBDs.** *Top*, visualization of PCs 1-3 for apo and glu-bound GluK2. The red arrow represents an axis of rotation for the LBD lower lobe. PC1 is a hinge bending motion for both apo and glu-bound. PC2 for the apo LBD is a 'sweeping' motion, whereas for glutamate-bound LBD it is a 'rocking' motion. PC3 for the apo form is the 'rocking' motion and for the glutamate-bound form PC3 corresponds to the 'sweeping' motion. *Bottom*, cumulative proportion of variance contained within PCs 1-20. Notably, 87% of variance is contained with PC1 for apo GluK2, but only 28.1% for glu-bound GluK2.

Fig. 3-11: Electrostatic potential surface of GluK2 and GluA2. Electrostatic potential surface of GluK2 and GluA2. Separate views of upper and lower LBD lobes: from the front (*left* and toward the binding cleft (*right*). *Top*, GluK2. *Bottom*, GluA2. In both GluA2 and GluK2 the glutamate-binding pocket can be seen in the areas of deep blue coloring. Qualitatively, GluA2 has more blue coloring than GluK2, which is also reflected in the total LBD charges (GluK2: -2, GluA2: +5)

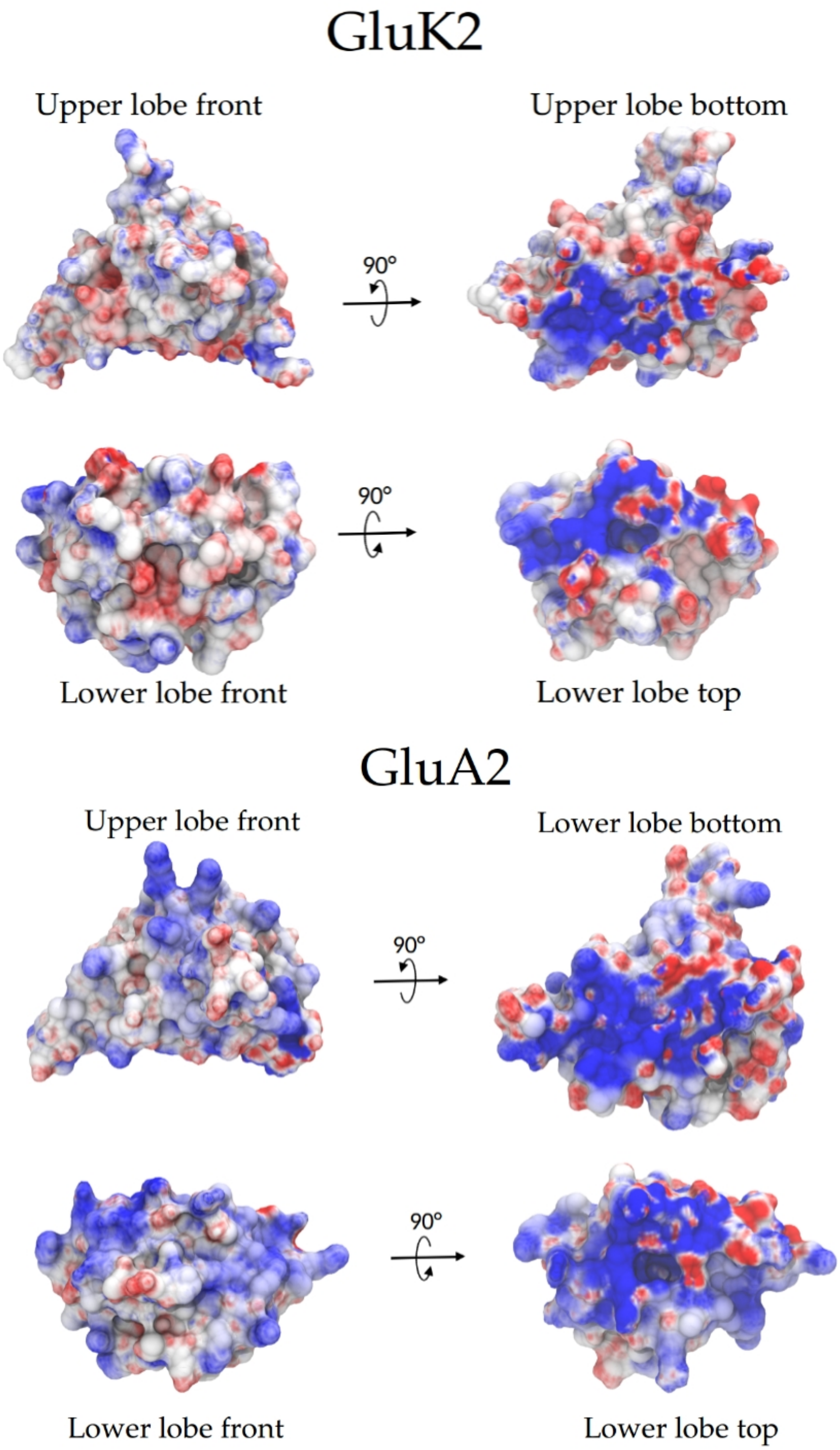

GluK2
Upper lobe front
Upper lobe bottom
90°
90°
Lower lobe front
Lower lobe top
GluA2
Upper lobe front
Lower lobe bottom
90°
90°
Lower lobe front
Lower lobe top

Fig. 3-11

Fig. 3-12: Electrostatic potential surface of GluN1, GluN2A, and GluN3A. Separate views of upper and lower LBD lobes: from the front (*left*) and toward the binding cleft (*right*). Both GluN1 (*top, left*) and GluN2A (*top, left*) have strong positive patches within the binding cleft, whereas GluN3A (*bottom*) has strong negative patches spread throughout the entire protein surface, with only a small positive patch at the glutamate binding site. These results are also reflected in the total charges of the LBDs: GluN1: +2, GluN2A: +1, GluN3A: -11

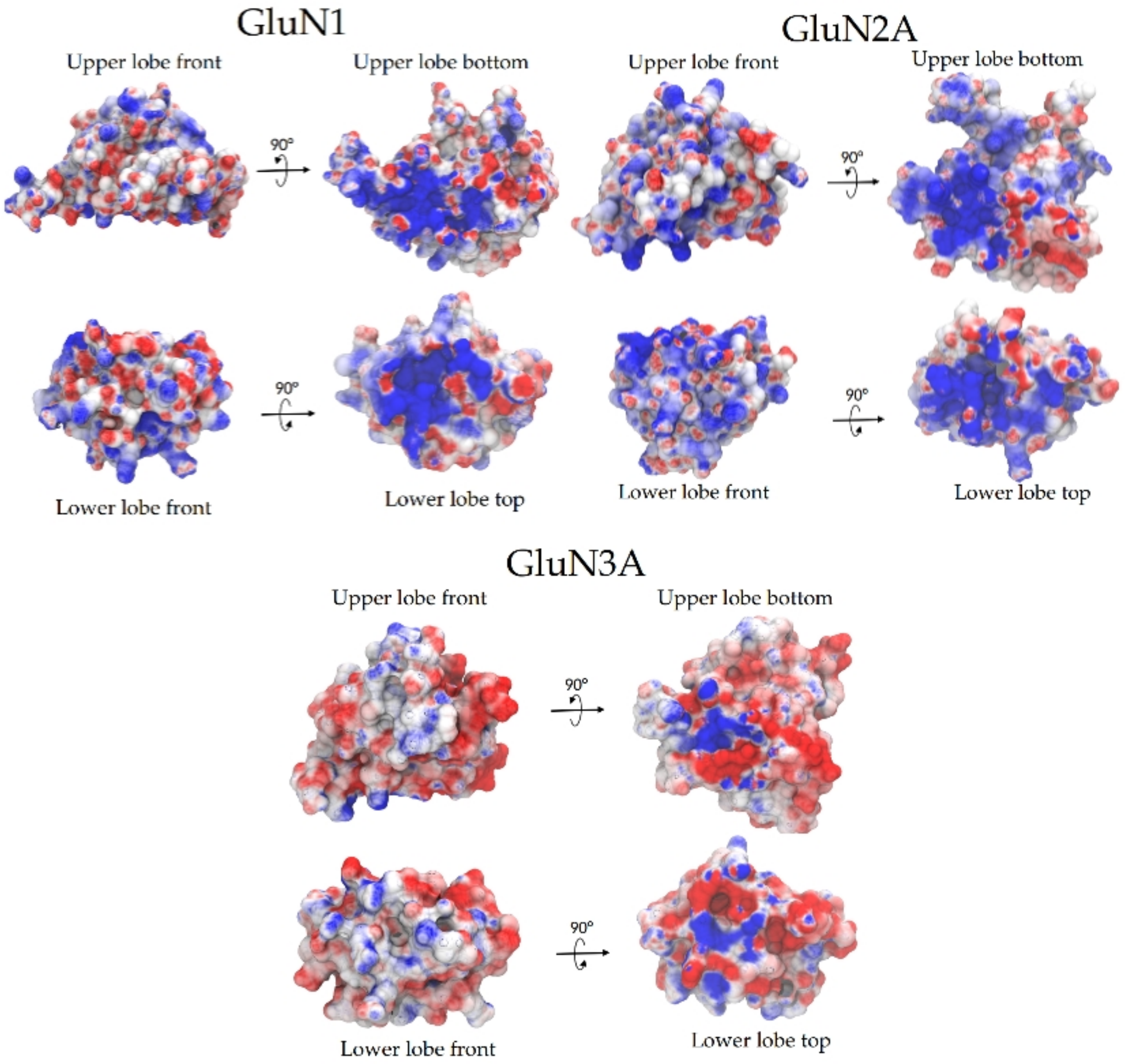

GluN1
Upper lobe front
Upper lobe bottom
90°
Lower lobe front
Lower lobe top
90°
GluN2A
Upper lobe front
Upper lobe bottom
90°
Lower lobe front
Lower lobe top
90°
GluN3A
Upper lobe front
Upper lobe bottom
90°
Lower lobe front
Lower lobe top
90°

Fig. 3-12

GluK2 LBD	ΔG_{conf} (kcal/mol)
WT Apo	-0.15
WT Glu	2.62
D656S Apo	-0.75
D656S Glu	1.89
D462A Apo	-0.16
D462A Glu	0.23

GluK2 LBD	$\Delta\Delta G_{\mathrm{conf}}$ (kcal/mol)
WT	2.77
D656S	2.64
D462A	0.39

Table 3.1: **Values of ΔG_{conf} and $\Delta\Delta G_{\mathrm{conf}}$** of GluK2 LBDs

References

[1] Ozawa, S., Kamiya, H., and Tsuzuki, K. Glutamate receptors in the mammalian central nervous system. *Progress in neurobiology*, 54(5):581–618, 1998.

[2] Kumar, J. and Mayer, M. L. Functional insights from glutamate receptor ion channel structures. *Annual review of physiology*, 75:313–337, 2013.

[3] Yuzaki, M. The $\delta 2$ glutamate receptor: 10 years later. *Neuroscience research*, 46(1):11–22, 2003.

[4] Traynelis, S. F., Wollmuth, L. P., McBain, C. J., Menniti, F. S., Vance, K. M., Ogden, K. K., Hansen, K. B., Yuan, H., Myers, S. J., and Dingledine, R. Glutamate receptor ion channels: structure, regulation, and function. *Pharmacological reviews*, 62(3):405–496, 2010.

[5] Pinheiro, P. S. and Mulle, C. Presynaptic glutamate receptors: physiological functions and mechanisms of action. *Nature Reviews Neuroscience*, 9(6):423, 2008.

[6] Contractor, A., Mulle, C., and Swanson, G. T. Kainate receptors coming of age: milestones of two decades of research. *Trends in neurosciences*, 34(3):154–163, 2011.

[7] Mayer, M. L. Glutamate receptors at atomic resolution. *Nature*, 440(7083):456, 2006.

[8] Gill, M. B., Vivithanaporn, P., and Swanson, G. T. Glutamate binding and conformational flexibility of ligand-binding domains are critical early determinants of efficient kainate receptor biogenesis. *Journal of Biological Chemistry*, 284(21):14503–14512, 2009.

[9] Lau, A. Y. and Roux, B. The free energy landscapes governing conformational changes in a glutamate receptor ligand-binding domain. *Structure*, 15(10):1203–1214, 2007.

[10] Lau, A. Y. and Roux, B. The hidden energetics of ligand binding and activation in a glutamate receptor. *Nature structural & molecular biology*, 18(3):283, 2011.

[11] Yao, Y., Belcher, J., Berger, A. J., Mayer, M. L., and Lau, A. Y. Conformational analysis of NMDA receptor GluN1, GluN2, and GluN3 ligand-binding domains reveals subtype-specific characteristics. *Structure*, 21(10):1788–1799, 2013.

[12] Dai, J. and Zhou, H.-X. Reduced curvature of ligand-binding domain free-energy surface underlies partial agonism at NMDA receptors. *Structure*, 23(1):228–236, 2015.

[13] Numano, R., Szobota, S., Lau, A. Y., Gorostiza, P., Volgraf, M., Roux, B., Trauner, D., and Isacoff, E. Y. Nanosculpting reversed wavelength sensitivity into a photoswitchable iGluR. *Proceedings of the National Academy of Sciences*, 106(16):6814–6819, 2009.

[14] Guo, Y., Wolter, T., Kubař, T., Sumser, M., Trauner, D., and Elstner, M. Molecular dynamics investigation of gluazo, a photo-switchable ligand for the glutamate receptor GluK2. *PloS one*, 10(8):e0135399, 2015.

[15] Weston, M. C., Gertler, C., Mayer, M. L., and Rosenmund, C. Interdomain interactions in AMPA and kainate receptors regulate affinity for glutamate. *Journal of Neuroscience*, 26(29):7650–7658, 2006.

[16] Birdsey-Benson, A., Gill, A., Henderson, L. P., and Madden, D. R. Enhanced efficacy without further cleft closure: reevaluating twist as a source of agonist efficacy in AMPA receptors. *Journal of Neuroscience*, 30(4):1463–1470, 2010.

[17] Yu, A., Alberstein, R., Thomas, A., Zimmet, A., Grey, R., Mayer, M. L., and Lau, A. Y. Molecular lock regulates binding of glycine to a primitive NMDA receptor. *Proceedings of the National Academy of Sciences*, 113(44):E6786–E6795, 2016.

[18] Chovancova, E., Pavelka, A., Benes, P., Strnad, O., Brezovsky, J., Kozlikova, B., Gora, A., Sustr, V., Klvana, M., Medek, P., et al. CAVER 3.0: a tool

for the analysis of transport pathways in dynamic protein structures. *PLoS computational biology*, 8(10):e1002708, 2012.

[19] Counterman, A. E. and Clemmer, D. E. Volumes of individual amino acid residues in gas-phase peptide ions. *Journal of the American Chemical Society*, 121(16):4031–4039, 1999.

[20] Svergun, D., Richard, S., Koch, M., Sayers, Z., Kuprin, S., and Zaccai, G. Protein hydration in solution: experimental observation by X-ray and neutron scattering. *Proceedings of the National Academy of Sciences*, 95(5):2267–2272, 1998.

[21] Mayer, M. L. Crystal structures of the GluR5 and GluR6 ligand binding cores: molecular mechanisms underlying kainate receptor selectivity. *Neuron*, 45(4):539–552, 2005.

[22] Vestergaard, B. Analysis of biostructural changes, dynamics, and interactions–small-angle X-ray scattering to the rescue. *Archives of biochemistry and biophysics*, 602:69–79, 2016.

[23] Alberstein, R., Grey, R., Zimmet, A., Simmons, D. K., and Mayer, M. L. Glycine activated ion channel subunits encoded by ctenophore glutamate receptor genes. *Proceedings of the National Academy of Sciences*, 112(44):E6048–E6057, 2015.

[24] Mamonova, T., Yonkunas, M. J., and Kurnikova, M. G. Energetics of the cleft closing transition and the role of electrostatic interactions in conformational rearrangements of the glutamate receptor ligand binding domain. *Biochemistry*, 47(42):11077–11085, 2008.

[25] Dolino, D. M., Adariani, S. R., Shaikh, S. A., Jayaraman, V., and Sanabria, H. Conformational selection and submillisecond dynamics of the ligand-binding domain of the N-methyl-D-aspartate receptor. *Journal of Biological Chemistry*, pages jbc–M116, 2016.

[26] Brooks, B. R., Brooks III, C. L., Mackerell Jr, A. D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., et al.

CHARMM: the biomolecular simulation program. *Journal of computational chemistry*, 30(10):1545–1614, 2009.

[27] Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. Scalable molecular dynamics with NAMD. *Journal of computational chemistry*, 26(16):1781–1802, 2005.

[28] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of chemical physics*, 79(2):926–935, 1983.

[29] Webb, B. and Sali, A. Comparative protein structure modeling using MODELLER. *Current protocols in bioinformatics*, 54(1):5–6, 2014.

[30] Feller, S. E., Zhang, Y., Pastor, R. W., and Brooks, B. R. Constant pressure molecular dynamics simulation: the Langevin piston method. *The Journal of chemical physics*, 103(11):4613–4621, 1995.

[31] Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. A smooth particle mesh Ewald method. *The Journal of chemical physics*, 103(19):8577–8593, 1995.

[32] Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., and Kollman, P. A. The weighted histogram analysis method for free-energy calculations on biomolecules. *Journal of computational chemistry*, 13(8):1011–1021, 1992.

[33] Souaille, M. and Roux, B. Extension to the weighted histogram analysis method: combining umbrella sampling with free energy calculations. *Computer physics communications*, 135(1):40–57, 2001.

[34] Grossfield, A. and Zuckerman, D. M. Quantifying uncertainty and sampling quality in biomolecular simulations. *Annual reports in computational chemistry*, 5:23–48, 2009.

[35] Svergun, D., Barberato, C., and Koch, M. H. CRYSOL–a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *Journal of applied crystallography*, 28(6):768–773, 1995.

[36] Grant, B. J., Rodrigues, A. P., ElSawy, K. M., McCammon, J. A., and Caves, L. S. Bio3d: an R package for the comparative analysis of protein structures. *Bioinformatics*, 22(21):2695–2696, 2006.

Chapter 4 - Conformational dynamics of a metabotropic glutamate receptor ligand binding domain

Reproduced in part from a manuscript *in prep*:

by

Tyler Wied and Albert Y. Lau

Metabotropic glutamate receptors (mGluRs) are dimeric class C GPCRs that mediate cellular responses to the neurotransmitter glutamate in the central nervous system. Crystal structures and single-molecule FRET results suggest that closure of the ligand-binding domains (LBDs) as well as reorientation of the dimer interface are necessary conformational transitions for receptor activation. The energetic factors that drive these coupled processes, however, are poorly understood. Here, we report free energy landscapes, or potentials of mean force (PMFs), for monomer mGluR3 LBDs in apo and glutamate-bound states as well as glutamate-bound dimer mGluR3 in the active and inactive conformations. PMFs of LBD monomers suggest the glu-bound LBD are bistable: the closed and open-cleft conformations are accessible for both glutamate-bound and apo LBDs, but with the apo closed-cleft conformation destabilized relative to the glutamate-bound. PMFs for the glu-bound dimers indicate a strong preference for the closed state. The energetics corresponding to the reorientation of the dimer interface are also examined.

4.1 Background

Class C GPCRs are characterized by obligate dimerization and large extracellular ligand-binding domains (LBDs). This class consists of the GABA$_\mathrm{B}$ receptors, the taste receptors for umami and sweet, metabotropic glutamate (mGluR), and calcium sensing receptors receptors [1]. The mGluRs are widely expressed in the mammalian central nervous system where they tune synaptic strength by modulating NMDA receptor activity [2]. Proper mGluR function is required for several higher-order cognitive processes, such as learning, memory, and synaptic plasticity. They also have been implicated neurological disorders like schizophrenia and autism [3,4]. The eight mGluRs are clustered into three phylogenetic groups: Group I (mGluR1, mGluR5), Group II (mGluR2-3), and Group III (mGluR4, mGluR6-8) [5].

Each mGluR subunit encodes an N-terminal clamshell ligand-binding domain (LBD) [6], a cysteine-rich domain (CRD), 7-pass transmembrane domain (TMD), and an intracellular C-terminal domain that serves as a scaffold for binding partners [7]. Closure of the LBD clamshell upon agonist binding is associated with a reorientation of the dimer interface which facilitates receptor activation [8]. smFRET studies reveal that isolated LBD dimers preserve the relevant conformational transitions between active and inactive state [9, 10]. Notably, isolated LBD dimers rapidly oscillate between the active and inactive states on a sub-millisecond timescale, where agonist increases the probability that the receptor visits the active state [11]. There is also a transient conformational intermediate between active and inactive poses [10]. Revealing the energetics and molecular mechanisms underlying these dynamics are essential for a full description of receptor activation.

Here, we present the potential of mean force (PMF) for mGluR3 LBDs. The LBD PMFs are computed using an umbrella sampling ([12]) scheme along two order parameters that describe the openness of the clamshell LBD. We do this for both glutamate-bound and apo mGluR3 monomers. Next, we compute PMFs for LBDs in both the active and inactive dimer poses. We find a unique stabilization of the closed state for LBDs in the presence of their dimer partners, and find altered dynamics in a lower lobe helix that may explain these observations. Finally, we present a PMF computed along a pathway corresponding to the dimer interface transition, which we describe with a dihedral angle Θ.

4.2 Results

4.2.1 Monomer mGluR3 PMFs

To examine the mechanics of mGluR3 LBD cleft motions, we compute the potential of mean force (PMFs) using umbrella sampling ([12]) along two order parameters that capture opening and closing of the LBD. The order parameters, ξ_1 and ξ_2, capture motion along each side of the LBD cleft **Fig. 4-1**. A similar approach to compute PMFs has been used in ionotropic glutamate receptors (iGluRs) LBDs ([13–15]), but has not been applied to clamshell-like LBDs in non-iGluR systems. PMFs were computed for monomeric glutamate-bound and apo mGluR3 (**Fig. 4-2**). Glutamate-bound mGluR3 is bistable, with free energy minima in the closed state at ((ξ_1, ξ_2) = (13.9, 16.3 Å); $W(\xi_1, \xi_2) = 0.174$ kcal/mol) and in the open state at ((ξ_1, ξ_2) = (28.2, 28.3 Å); $W(\xi_1, \xi_2) = 0.0$ kcal/mol). The closed-state local minimum is narrower than than the broad open state basin. There is a ∼3 kcal/mol free energy barrier near (ξ_1, ξ_2 = 12.5, 19 Å) separating open and closed conformations. Apo mGluR3 has a global free energy minimum at (ξ_1, ξ_2 = 20.8, 22.1 Å), which is located in a similar position to crystal structures of open mGluR3 LBDs bound to antagonists (see PDB: 3SM9). Error analysis for the monomer mGluR3 LBDs is shown in **Fig. 4-3**.

It would be computationally convenient in future mGluR simulations that examine additional degrees of freedom to reduce the description of the LBD state from 2D to 1D. To test whether a simulation in 1D accurately captures the results observed in 2D, we ran independent simulations of mGluR3 LBDs along a single order parameter, ξ_{12}, where $\xi_{12} = (\xi_1 + \xi_2)/2$. Qualitatively, the 1D PMFs agree with the 2D results: Glutamate-bound mGluR has a free energy minimum at $\xi_{12} = 15.2$ Å near the closed state, and another local minimum more open conformations near $\xi_{12} = 19$-21 Å with a ∼4 kcal/mol free energy barrier between the two minima (**Fig. 4-4**). The apo state has a global free energy minima at $\xi_{12} = 23.75$ Å, and is unstable in the closed state. However, there are differences between these 1D PMFs and the 1D projections shown in **Fig. 4-2**. For glutamate-bound mGluR, the open state minimum observed in the 2D PMFs near ∼28 Å is absent. In apo mGluR3, the closed state is substantially destabilized compared to the 2D simulations, $W(\xi_1, \xi_2) =\sim 1.0$ kcal/mol) in 2D, and $W(\xi_{12}) =\sim 4.5$ kcal/mol in 1D.

4.2.2 Dimer mGluR3 PMFs

Next, we computed PMFs for one mGluR3 LBD in the presence of its dimer partner, both in the inactive state and active state. The active and inactive dimer poses are each associated with a particular structure **Fig. 4-5**. The inactive state has a dimer interface between the LBD upper lobes (**Fig. 4-5**, *right*), and the transition to the active state is characterized by the formation of a dimer interface between the lower LBD lobes (**Fig. 4-5**, *left*). For the PMFs of mGluR dimer LBDs, the order parameters here are identical to those presented in the previous section with mGluR3 monomers, the only difference being that now these calculations are performed with the dimer partner present. For these calculations, both subunits are bound to glutamate, and the dimer partner not subject to umbrella sampling is constrained to a closed position (ξ_1, ξ_2 = 13.0, 16.5 Å). In the active state, there is a global free energy minimum at (ξ_1, ξ_2 = 13.0, 16.5 Å), and in the inactive state the global free energy minima is at (ξ_1, ξ_2 = 13.7, 16.5 Å) **Fig. 4-6**. Both PMFs for the active and inactive states possess one free energy minimum, with monotonically increasing free energy away from the free energy minima. The active state, however, has a much steeper free energy gradient compared to the inactive state which are readily seen in the 1D projections in **Fig. 4-6** *right*. Error analysis for dimer calculations are shown in **Fig. 4-7**.

Interestingly, glutamate-bound monomer LBD is bistable, but the glutamate-bound dimers are not. In the active state, this makes sense: extensive contacts between lobes mutually prevent LBD opening. For inactive state dimer, however, it is less clear why this is the case as there are no obvious steric blocks that prevent one LBD opening much like in the monomer. This suggests cooperativity between subunits which as been recently observed in vitro [16], and suggests some mechanism of allostery between subunits.

4.2.3 Free energy change between open and closed states

Calculating a ΔG between open and closed states would be a convenient way to summarize differences between PMFs. To do this, we differentiate between open and closed states by defining a pair of residues on the top and bottom lobes that form a bottleneck along a tunnel between bulk solvent and the binding pocket. These residues would restrict entry of glutamate to the binding pocket if the distance between the bottleneck residues is smaller than some threshold value. The LBD is defined as closed if the distance between the bottleneck residues is smaller than this value, and open if the distance is greater than this value. The CAVER program was used to identify bottleneck residue in mGluR3 [17]. In mGluR, the bottleneck was defined as the distance between C_γ of arginine 40 and C_β of serine 254. In a given conformation of mGluR, if the bottleneck distance is less than 12 Å, the LBD is closed, and open if the bottleneck distance is greater than or equal to 12 Å. 12 Å was chosen based on the geometry of a glutamate molecule, in accordance with previous results [15].

The relative probability of an mGluR3 LBD occupying an open or closed state is determined by the ratio of the partition functions,

$$\frac{\Omega_{\text{Open}}}{\Omega_{\text{Closed}}} = \frac{\int_{\text{Open}} e^{-W(\xi_1,\xi_2)/k_B T} d\xi_1 d\xi_2}{\int_{\text{Closed}} e^{-W(\xi_1,\xi_2)/k_B T} d\xi_1 d\xi_2} \tag{4.1}$$

where $W(\xi_1, \xi_2)$ is the PMF, and the region of integration corresponds to either (ξ_1, ξ_2) in which the bottleneck distance ≥12 Å (open) or <12 Å (closed); k_B is Boltzmann's constant, and T is temperature. The conformational free energy difference between open and closed states is then given by,

$$\Delta G_{\text{conf}} = -k_B T \ln\left(\frac{\Omega_{\text{open}}}{\Omega_{\text{Closed}}}\right) \tag{4.2}$$

The difference in free energy change between LBD systems is then given by,

$$\Delta\Delta G_{\text{conf}} = \Delta G^{\text{conf}}_{\text{glu}} - \Delta G^{\text{conf}}_{\text{apo}} \tag{4.3}$$

In monomer mGluR3, the apo state has a ΔG_{conf} of -1.19 kcal/mol, and the glutamate-bound LBD has a ΔG_{conf} of -0.60 kcal/mol. This corresponds to a stabilization of the closed state in the glutamate-bound LBD, $\Delta\Delta G_{\text{conf}}$, of 0.59 kcal/mol, which is on par with thermal energy. In the dimer system, the active state has a ΔG_{conf} of 10.31 kcal/mol and in the inactive state ΔG_{conf} is 6.99 kcal/mol. The closed state is stabilized in the active state by 3.32 kcal/mol compared to the inactive state. Dimer closed states are significantly stabilized relative to monomer mGluR3, 7.59 kcal/mol for the inactive state and 10.91 kcal/mol for the active state. Plots of how bottleneck distance scales with (ξ_1, ξ_2) for monomer and dimer mGluR3 LBDs are shown in **Fig. 4-8**. Free energies are summarized in **Table 4.1**.

4.2.4 Allostery in mGluR dimers

Why are the PMFs for glutamate-bound monomer mGluR3 LBDs different from the glutamate-bound dimer mGluR3s? In the active pose, there is a straightforward physical interpretation: the extensive contact between the lower lobes mutually prevent cleft opening in each subunit. However, in the inactive pose, the lower lobes are well separated. There are no obvious steric blocks to prevent LBD opening. One possibility is allosteric communication between subunits. Allostery involves communication between distal regions of a protein or protein complex. Here, we perform allosteric network analysis for the glutamate-bound monomer mGluR3 and inactive mGluR3 dimer. Briefly, proteins are represented as networks of nodes, which individual amino acids, and edges between nodes, which represent contact between residues. Here, the C_α atom of each residue is a node, and an edge is drawn between two nodes i, j if any two atoms belonging to residue i is within 5 Å of any atom in residue j for at least 75% of the simulation.

To get a quantify for how the dynamics of the dimer and monomer differ on the residue-level, covariance matrices, C_{ij}, were computed for both the dimer and monomer. This calculation will identify patches of residues whose dynamics change relative to the rest of the protein from one system to another. For the dimer, the covariance matrix is only generated for the LBD from which the PMF is computed. Then a difference covariance matrix (i.e. $\Delta C_{ij} = C_{ij,dimer} - C_{ij,monomer}$) was computed to highlight regions of the protein where dynamics differ between the two systems. There are two clusters of residues where visible differences are observed, referred to as region 1 and 2 (**Fig. 4-9**). Region 1 roughly corresponds to residues 95-112, which is highlighted in red on the structural model in the bottom left of **Fig. 4-9**. This loop notably has a disulfide linkage between the two mGluR3 subunits in the dimer, which likely alters its dynamics relative to the monomer state. This loop is unlikely to contribute to differences in thermodynamics of LBD opening because it is far removed from the LBD cleft. Region 2 is smaller than region 1, and corresponds roughly to residues 200-210. These residues are located on a helix in the lower lobe of the LBD. This helix is situated on the side of the LBD that faces the other dimer subunit.

We are currently in the process of running more dimer simulations to test the thermodynamic effect of changing the state of the dimer partner, namely, in the apo/open states. Additionally, with these results we may identify allosteric networks that change from one state to another.

4.2.5 Free energy change of dimer reorientation

To compute the energetics associated with the transition from the inactive to active state of the LBD, we use a 1D dihedral order parameter (Θ) that describes the state of the dimer (**Fig. 4-10**, *top*). The initial configurations span from $\Theta_{\text{inactive}} = 35°$ to $\Theta_{\text{active}} = 95°$. Umbrella sampling along this order parameter was used to calculate a PMF (**Fig. 4-10**, *bottom*). The PMF at first glance is not smooth and contains large error bars, even after more than 5 ns of simulation time per window. The global free energy minimum is located at $\Theta = 99.5°$, which is near the edge of the Θ sampled with umbrella sampling and is also $\sim$2.5 kcal/mol lower in free energy than any other region of the PMF. Other than this, qualitatively, the PMF appears to have two minima separated by a free energy barrier. The minima are near $\Theta = 35°$ and $\Theta = 80°$, although this second minimum could properly extend from $\Theta = 70° - 100°$, given the large error bars.

4.3 Discussion

To probe the space of accessible conformations for an mGluR LBD, we computed 2D PMFs of apo and glutamate-bound mGluR3 monomer LBDs. Both monomer LBDs are highly conformationally flexible, with low free energy regions spanning open and closed conformations. The apo PMF has a slightly less accessible closed state than the glutamate-bound LBD, which exhibits two-state behavior. To quantify the difference between the PMFs, we classified conformations into open and closed states using a cutoff distance value between two bottleneck residues lining a binding tunnel. The closed state for apo mGluR3 is roughly 0.6 kcal/mol destabilized relative to glutamate-bound monomer. We also computed monomer mGluR3 PMFs in 1D, which would be useful in simulations that examine additional degrees of freedom. The 1D PMFs are similar to the 1D projections of 2D PMFs. However, differences are observed, such as the discrepancy in the position of the open-state minimum for glutamate bound mGluR3. This discrepancy is likely explained by the choice of initial configurations for umbrella sampling, which were generated along a diagonal of (ξ_1, ξ_2). This diagonal along which conformations were selected has a slope equal to one, but the slope of the 2D PMF corresponding to the pathway that minimizes the free energy is less than 1, so the initial configurations for open conformations in the 1D PMF are too far removed from the 2D minimum. This could be avoided in future studies by increasing the sampling time, which is currently insufficient (0.8 ns/window), or with a better choice of initial configurations. In apo mGluR3, the closed state is substantially destabilized in the 1D simulations compared to the 2D simulations. It's unclear why this discrepancy exists, but it possibly reflects normal variance in results (the two 1D plots are just outside of ~ 1 standard deviation of each other in the closed states).

We also computed 2D PMFs for mGluR3 LBDs in the presence of their dimer partners in active and inactive dimer states. The active state stabilizes closed cleft LBD states compared to the inactive state, which is likely explained by extensive dimer contacts between the upper and lower lobes in the active state, which prevent LBD cleft opening. For both the active and inactive states, the closed LBD state is significantly stabilized compared to monomer mGluR3. This makes sense in the active form, where the lower lobes likely mutually reinforce closed states. In the inactive state, however, their is nothing obviously blocking LBD opening. We found that a helix in the lower lobe changes its dynamics in the dimer states vs the monomer state. Notably, this helix is situated on the side of the LBD that faces the other dimer subunit. One intriguing possibility explaining this change is an allosteric signal from the dimer partner, although future will will have to test this. Cooperativity has been documented in mGluRs [16]. Importantly, the input data for this analysis is unweighted. Future work should appropriately Boltzmann-weight input data for a more accurate representation of dynamics. Additionally, it would be interesting to examine the change in dynamics in particular regions of the PMF, perhaps near the closed state. Future work should also examine the inactive state PMF in various states of the LBD dimer partner. especially apo and/or open states. To look at allostery, dynamical network analysis can be readily performed much like that done in [18], and is available as an analysis package in VMD.

To calculate the PMF along a pathway corresponding to the transition between inactive and active states of the dimer, we defined a 1D dihedral order parameter, Θ, to describe the state of the dimer. The resulting PMF is noisy with large errors. The source of the large error bars is unknown. The system has been extensively sampled (5 ns/window), but it is possible convergence will happen with more sampling. It is also possible an additional order parameter is required to sample this space. Generally there appears to be two regions of lower free energy separated by a free energy barrier. This system has been extensively sampled (5 ns/window) and does not appear to be closer to convergence than at 4 or 3 ns/window (data not shown). Future work here should focus on potentially revisiting the order parameters for this system, or perhaps adjusting the force constant.

4.4 Methods

4.4.1 Molecular dynamics simulation

MD simulations were carried out in the CHARMM [19] and NAMD [20] programs with TIP3 explicit water model [21]. PDB structure 2E4U was used as a representative mGluR 3 structure, with the CRD removed [22]. 2E4U is in an inactive pose. To model the active state, the 2E4U LBDs were structurally aligned to the mGluR3 LBDs in 3SM9, which is an active pose. Residues not modeled in the crystal structure were built using the MODELLER [23] program. The disulfide bridge between cysteine 103 of each subunit was built into our model [24]. Crystallographic waters buried within the LBD were maintained, whereas waters outside the LBD were deleted. Orthorhombic water boxes (monomer: $110 \times 80 \times 65$ Å^3, dimer-active: $116 \times 87 \times 87$ Å^3, and dimer-inactive $118 \times 98 \times 83$ Å^3) with 150 mM NaCl were used for all simulations, with an offset to balance protein charge and to set the total system charge to 0. Following equilibration in NVT ensemble, production runs were carried out in the NPT, with $P = 1$ atm and $T = 300$ K [25]. Long range electrostatic interactions were computed with the Ewald method [26]. In simulations with a glutamate ligand, the γ carboxyl group was restrained with a half-harmonic potential to the guanidine nitrogen atoms of arginine 44. The force constant to keep glutamate tethered to arginine 44 was set to 2 kcal/mol/Å if the ligand drifted more than 3.2 Å from the arginine 44.

4.4.2 Umbrella sampling

Umbrella sampling of LBD opening was implemented along 2 order parameters. The first order parameter, ξ_1, is the distance between the center-of-masses of amino acids 126-128 and 254-256, and the second, ξ_2 between the center-of-masses of amino acids 364-366 and 283-285. The 1D Umbrella sampling was done by performed by an order parameter between the center-of-masses of residues 126-128 and 364-366 in the top lobe and 254-256 and 283-285 in the bottom lobe. 190 umbrella windows were generated along the order-parameters for the 2D calculations and each was run for 2.2 ns/window in glutamate-bound mGluR3 monomer, and 1.4 ns/window for apo mGluR monomer. 20 windows were generated for the 1D calculations, and production runs were carried out for 0.8 ns/window for both apo and glutamate-bound mGluR3. Unbiasing and free energy calculations were carried out with the weighted histogram analysis method (WHAM) [27].

For simulations of of the dimer, umbrella sampling was performed with one LBD. It's dimer partner was restrained to a closed conformation, $((\xi_1, \xi_2) = (14,\ 17\ \text{Å}))$ with a force constant of 2 kcal/mol/Å. The windows for umbrella sampling are the same as in the monomer (above). Umbrella sampling was carried out for 1 ns/window in the active state, and 1.6 ns/window in the inactive state.

To compute the PMF along the pathway corresponding to the dimer interface transition between the inactive and active states, 70 intermediate states were generated with a linear interpolation morph. Umbrella sampling was conducted along a 1D order parameter corresponding to a pseudo-dihedral angle (Θ) describing the dimeric state. The dihedral angle is defined between 1) residues 84-87 of subunit A, 2) residues 134-137 of subunit A, 3) residues 134-137 of subunit B, and 4) residues 84-87 of subunit B. The initial configuration along this pathway ranges from $\Theta =$

$30^\circ - 100^\circ$. Force constants for the umbrella windows were set to 2 kcal/mol/deg^2. Each window was equilibrated for 3.0 ns, with 5.4 ns per window in production runes (108 ns total). PMFs were computed with WHAM.

4.4.3 Covariance matrix calculations

The covariance matrix is given by,

$$C_{ij} = \langle x_i - \bar{x}_i \rangle \langle x_j - \bar{x}_j \rangle \tag{4.4}$$

where $\bar{x}_i$ is the mean position of the ith atom and $\langle \cdot \rangle$ represents a mean value. For covariance matrix calculations of mGluR3, residues were represented by their C_α atoms. The covariance matrices were generated using the carma program [28]. Unweighted DCD files from umbrella sampling production runs were used as input into carma.

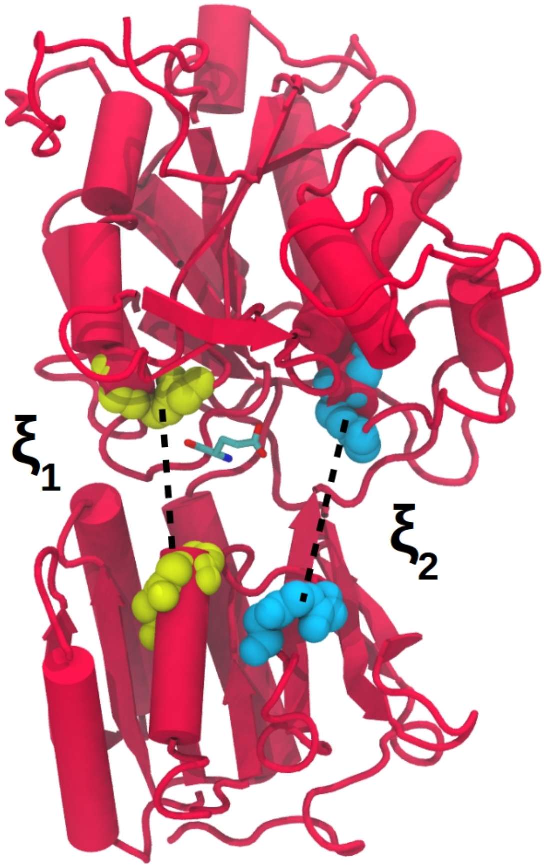


Fig. 4-1: **mGluR3 order parameters.** ξ_1 and ξ_2 are the 2D order parameters that describes the conformational state of mGluR LBD cleft in our simulations. ξ_1 and ξ_2 each indicate a distance, represented by dashed lines, between the residues shown in blue spheres.

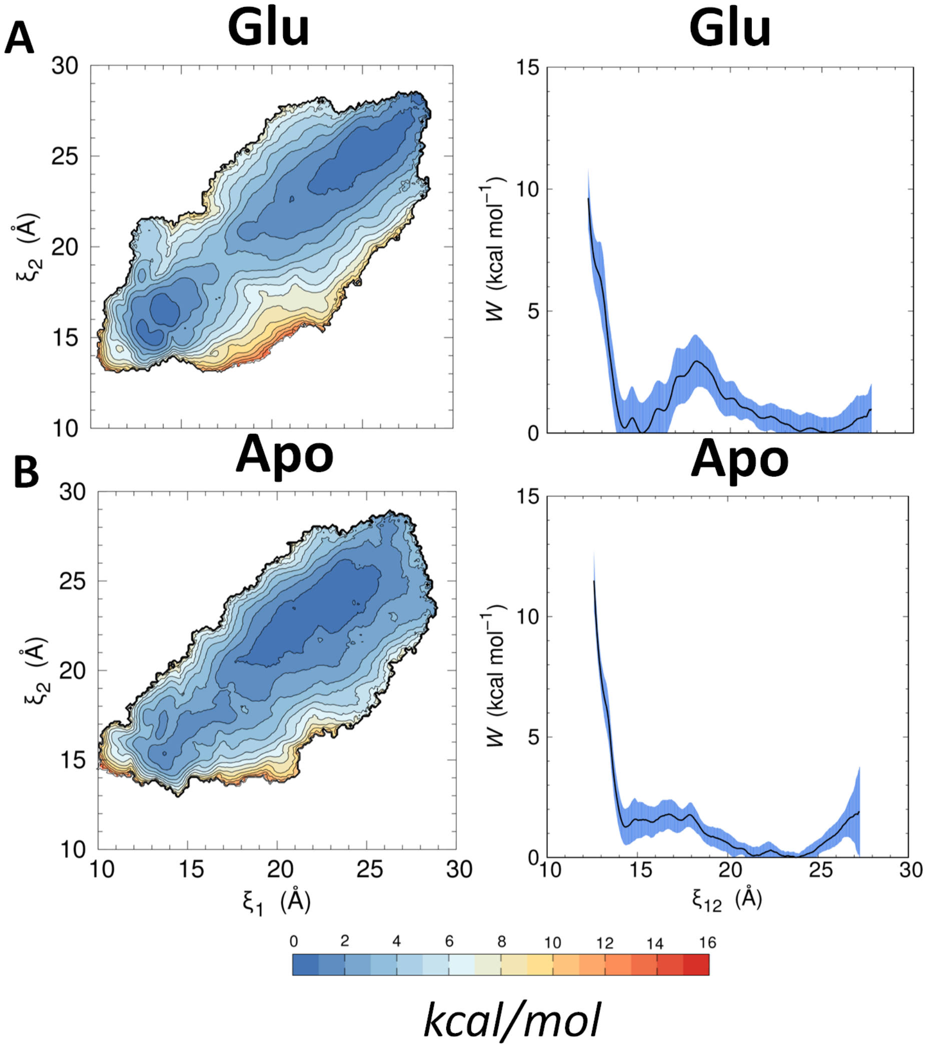


Fig. 4-2: **Potential of mean force for apo and glu-bound mGluR3 monomer.** **A**. Glutamate-bound, and **B**. apo. The 2D plots are *left*, and 1D projections are *right*, where the black line is the mean free energy value, and blue is the standard deviation. The 1D plots are computed along a hybrid order parameter ξ_{12}, where $\xi_{12} = (\xi_1 + \xi_2)/2$.

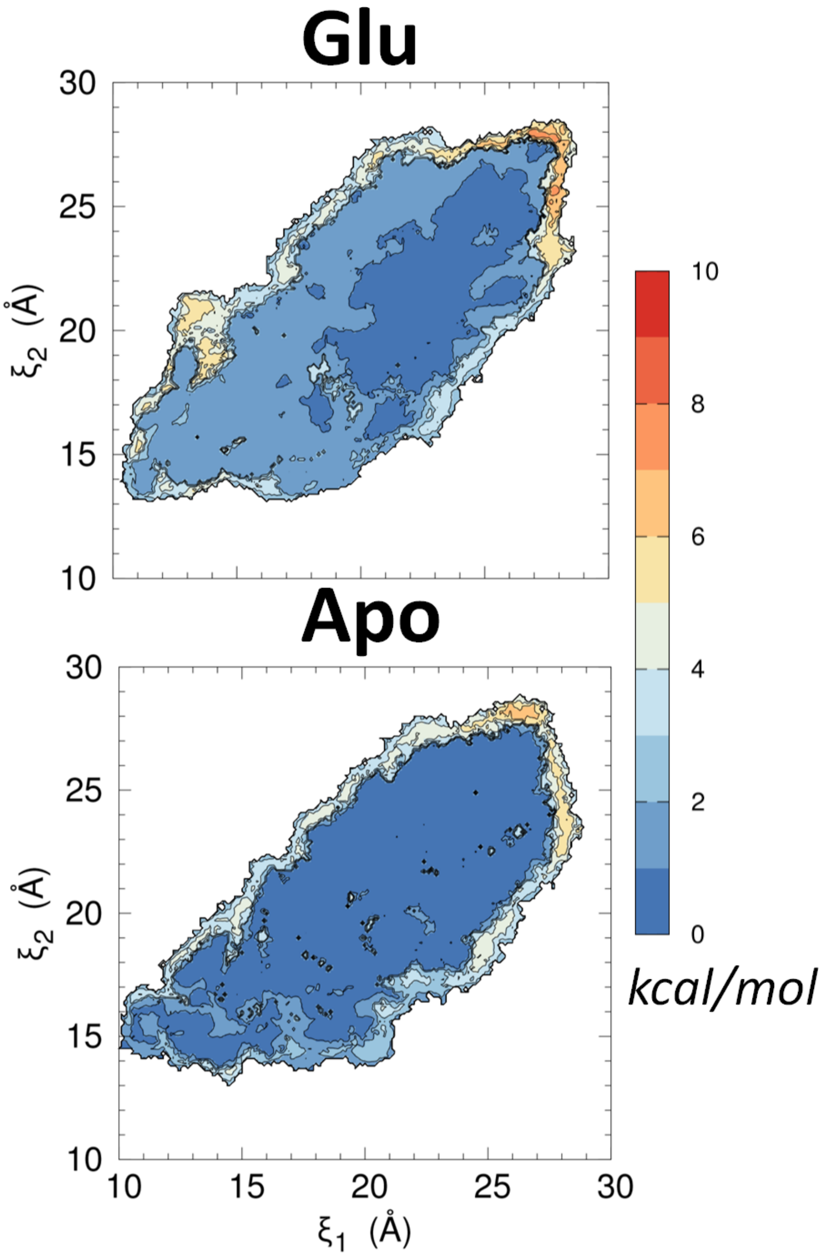


Fig. 4-3: **Standard deviation of monomer 2D PMFs.** The standard deviation of PMFs shown in **Fig. 4-2**, glutamate-bound mGluR3 is on *top*, and apo mGluR3 is on *bottom*.

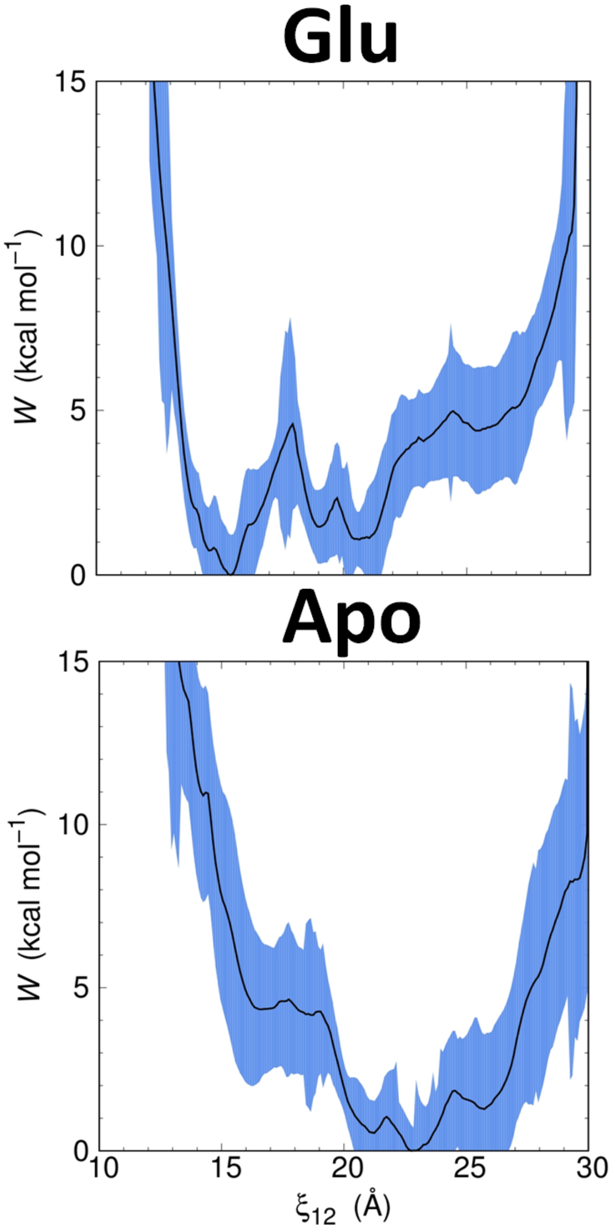


Fig. 4-4: **PMF of mGluR3 monomer LBD conducted with a 1D order parameter.** 1D PMFs of glutamate-bound *top*, and apo *bottom* mGluR3. The umbrella windows for these simulations were placed along ξ_{12}.

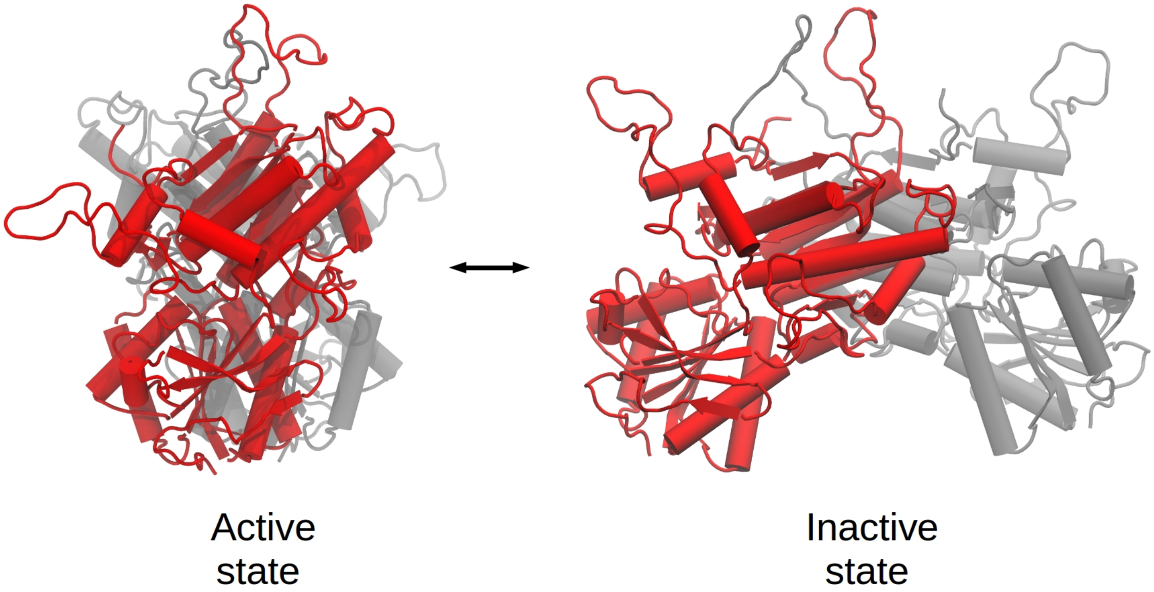


Fig. 4-5: **Structure of active and inactive mGluR3 dimer states.** Active state *left* and inactive state *right*.

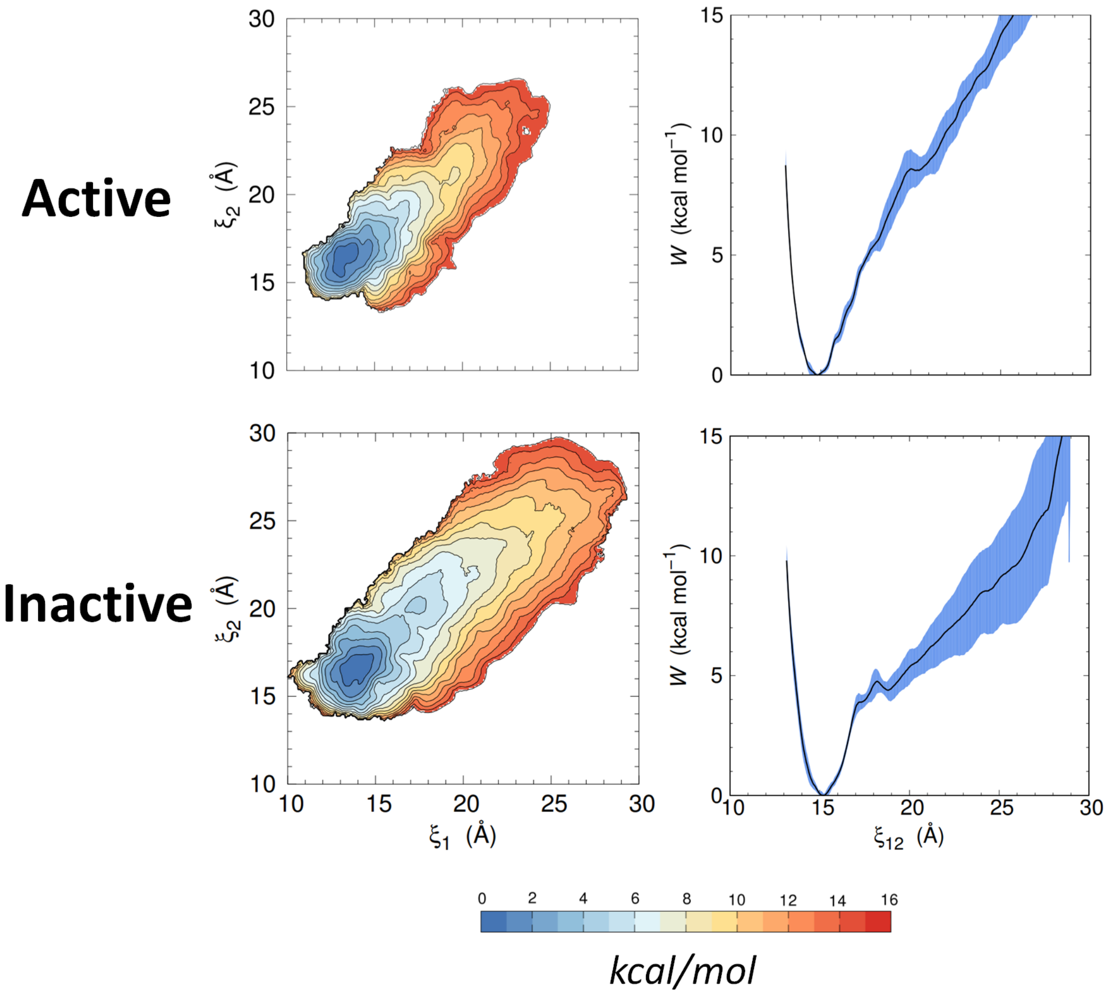


Fig. 4-6: **PMFs of mGluR3 LBD opening in dimer states.** *Top*, active state. *Bottom*, inactive state. 2D PMFs are shown on the left, and 1D projections along ξ_{12} are on the right.

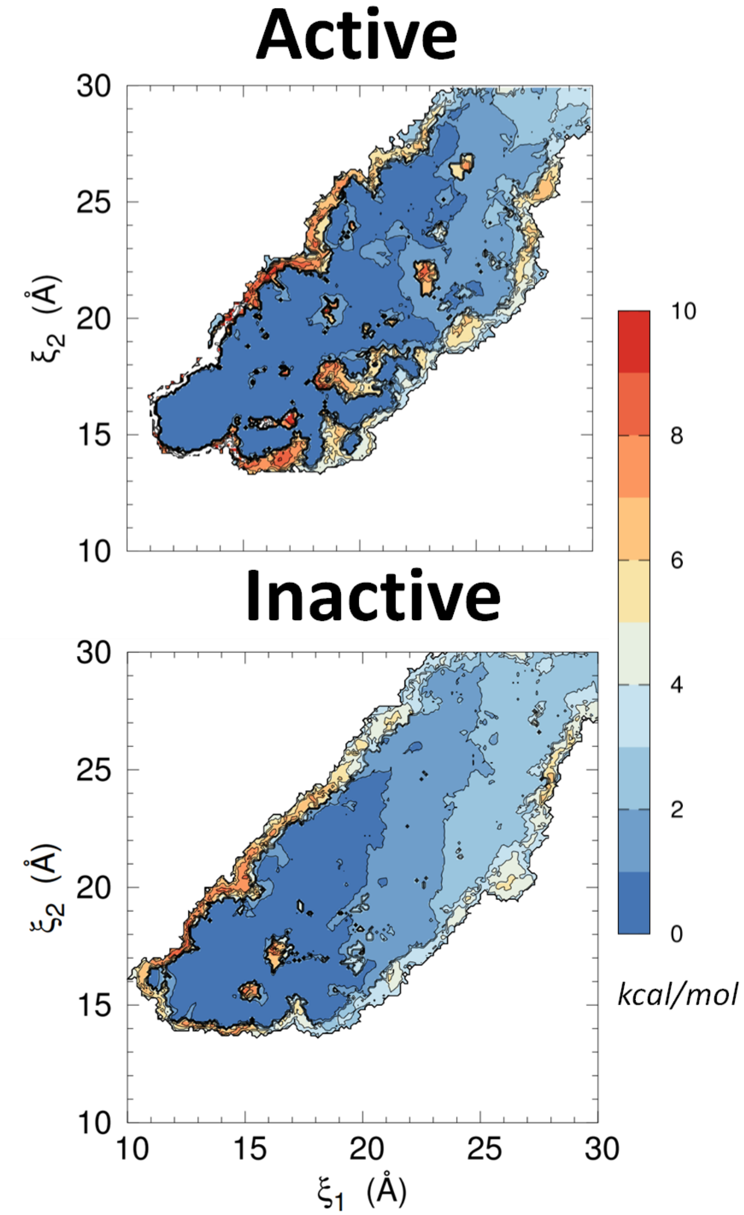


Fig. 4-7: **Standard deviation of dimer simulations.** Active (*top*) and inactive (*bottom*) mGluR dimers.

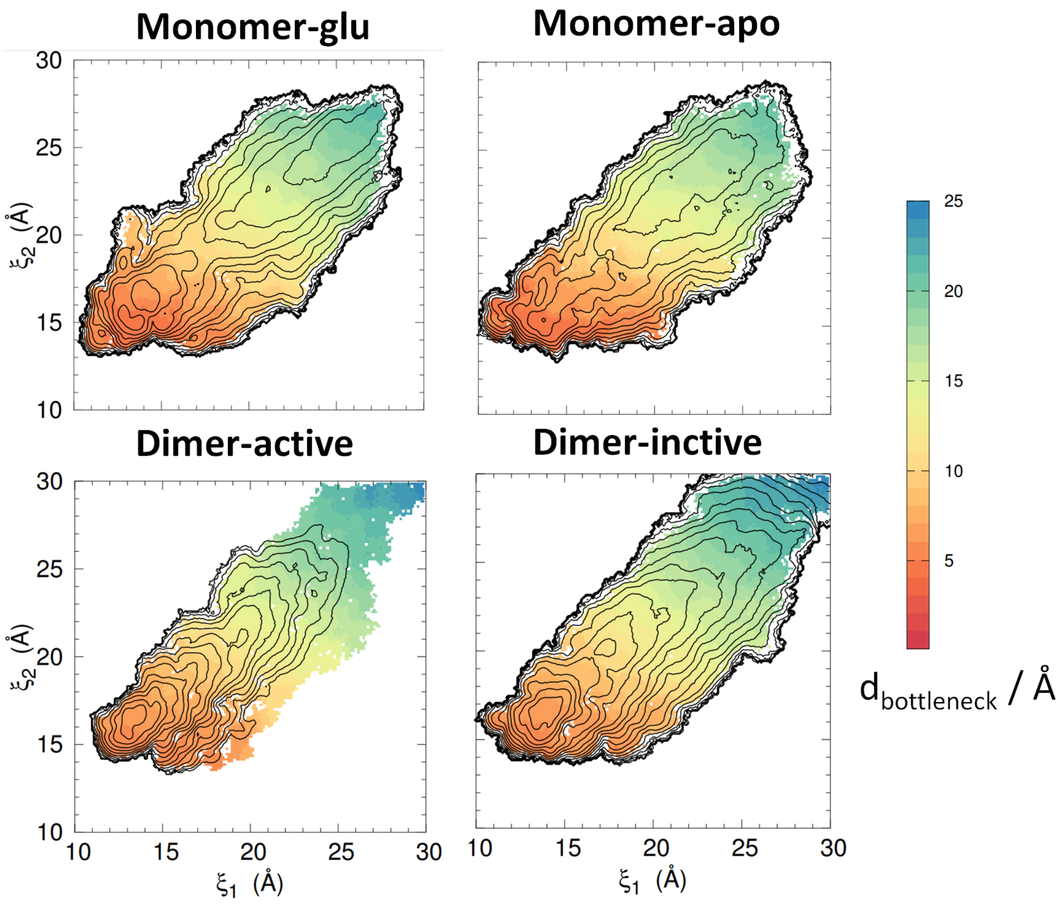


Fig. 4-8: **Distance between the residues defining the bottleneck in mGluR3.** The bottleneck distance is defined between C_γ of arginine 40 and C_β of serine 254. The contours correspond to the PMF and the color corresponds to the average distance between the residues in a given 0.1×0.1 Å^2 (ξ_1, ξ_2) window.

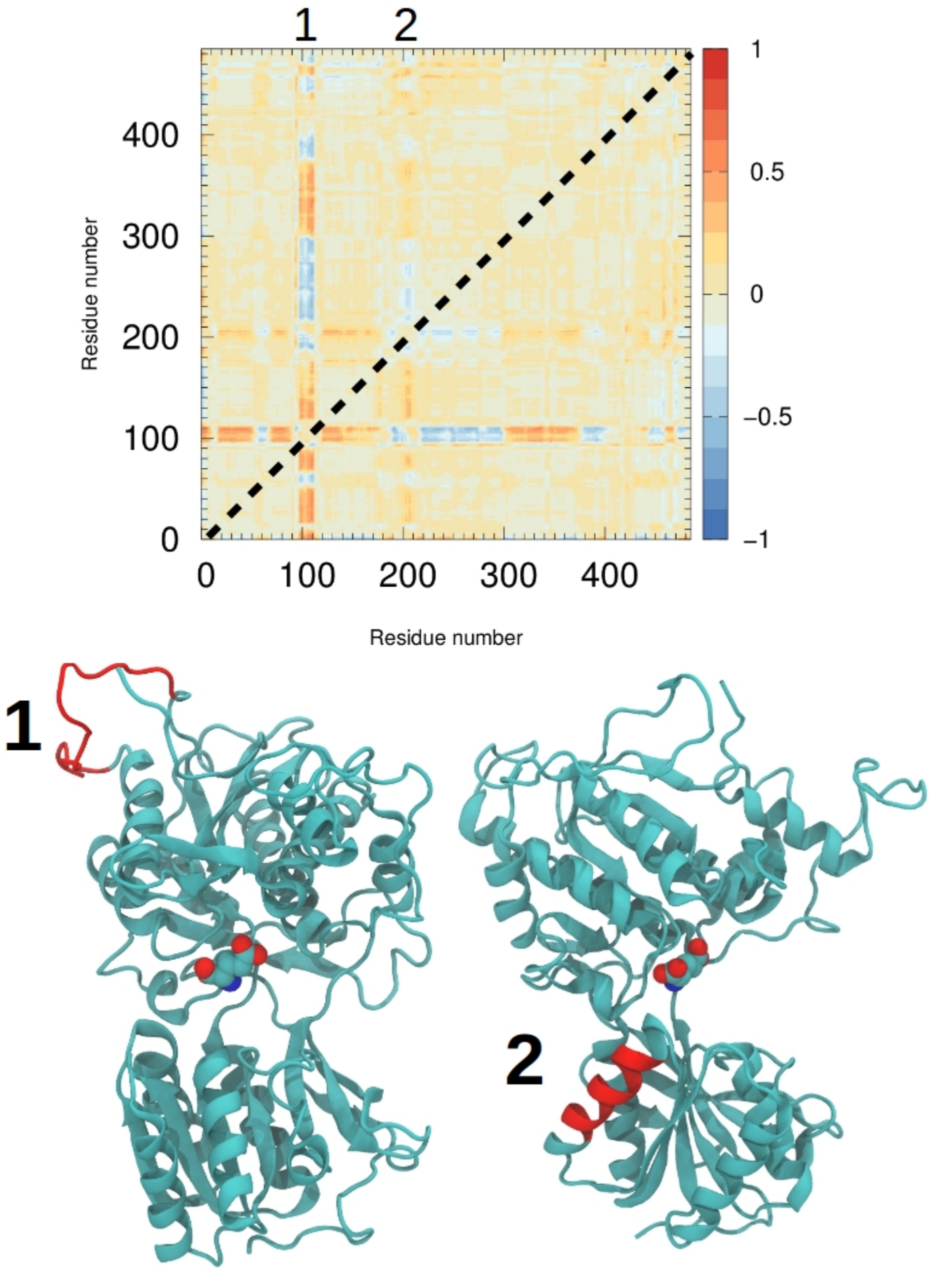


Fig. 4-9: **Dynamic differences between monomer and dimer mGluR3 LBDs.** *Top*, difference covariance matrix (ΔC_{ij}) between glutamate-bound inactive dimer and monomer LBDs. Regions with altered dynamics are labeled with 1 and 2. The matrix is symmetric over the dashed line. *Bottom,* structures of mGluR3 LBD with regions 1 and 2 with altered dynamics highlighted in red. Region 1, left, corresponds to a lop on the upper lobe. Region 2, right, corresponds to a helix in the lower lobe.

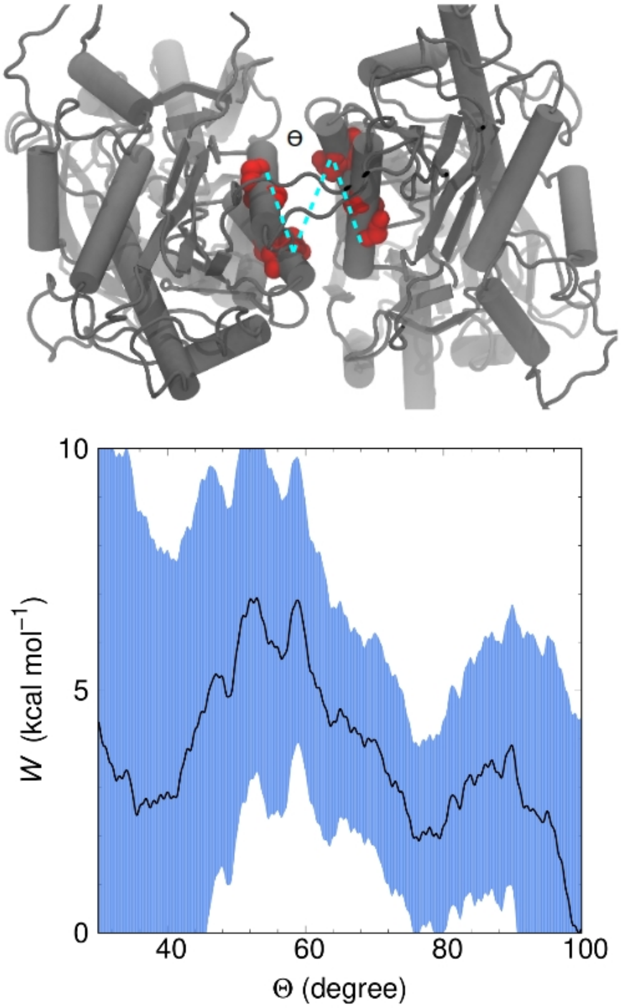


Fig. 4-10: **PMF of dimer interface reorientation.** *Top*, Atoms defining Θ are shown with red spheres. *Bottom*, PMF along Θ.

System	ΔG_{conf} (kcal/mol)
Monomer-Apo	-1.19
Monomer-Glu	-0.60
Dimer-Active	10.31
Dimer-Inactive	6.99

System	$\Delta\Delta G_{\text{conf}}$ (kcal/mol)
Monomer	0.59
Active - Inactive	3.32
Inactive - Monomer-Glu	7.59

Table 4.1: **Values of ΔG_{conf} and** $\Delta\Delta G_{conf}$ of mGluR3 LBDs.

References

[1] Møller, T. C., Moreno-Delgado, D., Pin, J.-P., and Kniazeff, J. Class C G protein-coupled receptors: reviving old couples with new partners. *Biophysics reports*, 3(4-6):57–63, 2017.

[2] Ambrosini, A., Bresciani, L., Fracchia, S., Brunello, N., and Racagni, G. Metabotropic glutamate receptors negatively coupled to adenylate cyclase inhibit N-methyl-D-aspartate receptor activity and prevent neurotoxicity in mesencephalic neurons in vitro. *Molecular pharmacology*, 47(5):1057–1064, 1995.

[3] Hammond, J., Shan, D., Meador-Woodruff, J., and McCullumsmith, R. Evidence of glutamatergic dysfunction in the pathophysiology of schizophrenia. In *Synaptic Stress and Pathogenesis of Neuropsychiatric Disorders*, pages 265–294. Springer, 2014.

[4] Chiocchetti, A. G., Bour, H. S., and Freitag, C. M. Glutamatergic candidate genes in autism spectrum disorder: an overview. *Journal of Neural Transmission*, 121(9):1081–1106, 2014.

[5] Niswender, C. M. and Conn, P. J. Metabotropic glutamate receptors: physiology, pharmacology, and disease. *Annual review of pharmacology and toxicology*, 50:295–322, 2010.

[6] Kunishima, N., Shimada, Y., Tsuji, Y., Sato, T., Yamamoto, M., Kumasaka, T., Nakanishi, S., Jingami, H., and Morikawa, K. Structural basis of glutamate recognition by a dimeric metabotropic glutamate receptor. *Nature*, 407(6807):971, 2000.

[7] Enz, R. Structure of metabotropic glutamate receptor c-terminal domains in contact with interacting proteins. *Frontiers in molecular neuroscience*, 5:52, 2012.

[8] Huang, S., Cao, J., Jiang, M., Labesse, G., Liu, J., Pin, J.-P., and Rondard, P. Interdomain movements in metabotropic glutamate receptor activation. *Proceedings of the National Academy of Sciences*, page 201107775, 2011.

[9] Doumazane, E., Scholler, P., Fabre, L., Zwier, J. M., Trinquet, E., Pin, J.-P., and Rondard, P. Illuminating the activation mechanisms and allosteric properties of metabotropic glutamate receptors. *Proceedings of the National Academy of Sciences*, page 201215615, 2013.

[10] Vafabakhsh, R., Levitz, J., and Isacoff, E. Y. Conformational dynamics of a class C G-protein-coupled receptor. *Nature*, 524(7566):497, 2015.

[11] Olofsson, L., Felekyan, S., Doumazane, E., Scholler, P., Fabre, L., Zwier, J. M., Rondard, P., Seidel, C. A., Pin, J.-P., and Margeat, E. Fine tuning of sub-millisecond conformational dynamics controls metabotropic glutamate receptors agonist efficacy. *Nature communications*, 5:5206, 2014.

[12] Torrie, G. M. and Valleau, J. P. Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *Journal of Computational Physics*, 23(2):187–199, 1977.

[13] Lau, A. Y. and Roux, B. The free energy landscapes governing conformational changes in a glutamate receptor ligand-binding domain. *Structure*, 15(10):1203–1214, 2007.

[14] Yao, Y., Belcher, J., Berger, A. J., Mayer, M. L., and Lau, A. Y. Conformational analysis of NMDA receptor GluN1, GluN2, and GluN3 ligand-binding domains reveals subtype-specific characteristics. *Structure*, 21(10):1788–1799, 2013.

[15] Yu, A., Alberstein, R., Thomas, A., Zimmet, A., Grey, R., Mayer, M. L., and Lau, A. Y. Molecular lock regulates binding of glycine to a primitive NMDA receptor. *Proceedings of the National Academy of Sciences*, 113(44):E6786–E6795, 2016.

[16] Levitz, J., Habrian, C., Bharill, S., Fu, Z., Vafabakhsh, R., and Isacoff, E. Y. Mechanism of assembly and cooperativity of homomeric and heteromeric metabotropic glutamate receptors. *Neuron*, 92(1):143–159, 2016.

[17] Chovancova, E., Pavelka, A., Benes, P., Strnad, O., Brezovsky, J., Kozlikova, B., Gora, A., Sustr, V., Klvana, M., Medek, P., et al. CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures. *PLoS computational biology*, 8(10):e1002708, 2012.

[18] Sethi, A., Eargle, J., Black, A. A., and Luthey-Schulten, Z. Dynamical networks in tRNA: protein complexes. *Proceedings of the National Academy of Sciences*, 106(16):6620–6625, 2009.

[19] Brooks, B. R., Brooks III, C. L., Mackerell Jr, A. D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., et al. CHARMM: the biomolecular simulation program. *Journal of computational chemistry*, 30(10):1545–1614, 2009.

[20] Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. Scalable molecular dynamics with NAMD. *Journal of computational chemistry*, 26(16):1781–1802, 2005.

[21] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of chemical physics*, 79(2):926–935, 1983.

[22] Muto, T., Tsuchiya, D., Morikawa, K., and Jingami, H. Structures of the extracellular regions of the group II/III metabotropic glutamate receptors. *Proceedings of the National Academy of Sciences*, 104(10):3759–3764, 2007.

[23] Webb, B. and Sali, A. Comparative protein structure modeling using MODELLER. *Current protocols in bioinformatics*, 54(1):5–6, 2014.

[24] Rondard, P., Liu, J., Huang, S., Malhaire, F., Vol, C., Pinault, A., Labesse, G., and Pin, J.-P. Coupling of agonist binding to effector domain activation in metabotropic glutamate-like receptors. *Journal of Biological Chemistry*, 2006.

[25] Feller, S. E., Zhang, Y., Pastor, R. W., and Brooks, B. R. Constant pressure molecular dynamics simulation: the Langevin piston method. *The Journal of chemical physics*, 103(11):4613–4621, 1995.

[26] Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. A smooth particle mesh Ewald method. *The Journal of chemical physics*, 103(19):8577–8593, 1995.

[27] Souaille, M. and Roux, B. Extension to the weighted histogram analysis method: combining umbrella sampling with free energy calculations. *Computer physics communications*, 135(1):40–57, 2001.

[28] Glykos, N. M. Software news and updates carma: A molecular dynamics analysis program. *Journal of computational chemistry*, 27(14):1765–1768, 2006.

Chapter 5 - Structural characterization of insect ionotropic receptor ligand binding domains

5.1 Background

Insects recognize environmental chemical stimuli with ion channel receptors [1]. The ionotropic receptors (IRs) are olfactory receptors expressed in chemosensory neurons in *Drosophila*, and are homologous to ionotropic glutamate receptors (iGluR) [2]. In addition to olfaction, some IRs regulate cold avoidance and humidity sensing ([3,4]). IRs are probably expressed as hetero tetrameric ion channels, much like iGluRs [5]. The subunits of the tetramer are composed of odor-specific subunits, which bind odorants, and co-receptor subunits, which are required for receptor assembly and trafficking [5,6].

To date, there are no atomic structures of an IR, or its isolated domains. Here, I describe the construct design and initial expression and purification attempts for the isolated LBDs of *Drosophila* IR8a, IR25a, IR75a, and IR84a in *E. coli* and a baculovirus insect cell expression system. In *E. coli* expression systems, IR84a and IR75a partition to insoluble inclusion bodies, even when tagged with a SUMO solubility tag [7]. Co-expression of an odor-specific subunit (IR75a) with a co-receptor (IR8a) was also unsuccessful. Various strategies were employed to re-fold LBDs from inclusion bodies: on-column refolding, step-wise and flash dilution, and dialysis (see [8]). Generally, re-folding was unsuccessful, but a few promising leads were observed in the presence of aggregation inhibitors. In insect cell systems, successful expression is observed for the IR25a LBD, although most protein partitions to inclusion bodies, but some protein partitions to the soluble fraction. Baculovirus expression looks promising, but more work is necessary. The purpose of this chapter is to serve as a roadmap of strategies that have been tried to obtain soluble IRs, and to highlight promising avenues forward.

5.2 Results

5.2.1 Construct design for IR LBD expression in *E. coli.*

For expression in *E. coli* systems, constructs were built for two co-receptors subunit LBDs, IR8a and IR25a, and two odor-specific subunit LBDs, IR84a and IR75a. The odor-specific subunits were chosen because there are validated ligands for these receptors. When expressed with the IR8a co-receptor, IR75a is activated by propionic acid and IR84a by phenylacetaldehyde [5]. LBD sequences were designed using isolated iGluR LBDs as a template (**Fig. 5-1**). The S1 and S2 segments were connected by a glycine-threonine linker analogous to what is typical of iGluR LBD constructs. IR75a was built with an N-terminal SUMO tag to improve solubility [7]. Sequences for each of the IR LBDs presented here are available in **Table 5.1**.

5.2.2 Expression and purification of IR84a LBD in *E. coli.*

IR84a was expressed in *E. coli* and purified from the insoluble fraction solubilized in the denaturant 6 M guanidine hydrochloride (GdnHCl) **Fig. 5-2** (see **section 3.4**). IR84a expresses exclusively in the insoluble fraction and solubilizes well in 6 M GdnHCl but not urea (data not shown). Denatured IR84a was purified on a nickel column and subjected to various refolding strategies. Re-folding on a nickel column was unsuccessful (**Fig. 5-3**): when GdnHCl was washed away, followed by an imidazole elution buffer, no detectable protein was collected. However, when again washed with a GdnHCl buffer, and then eluted with a buffer containing both imidazole and GdnHCl, significant protein is collected. These results suggest that IR84a forms aggregates on the column that are too large to elute. Perhaps this strategy would be more successful with more dilute protein.

Flash dilution and buffer exchange strategies were also unsuccessful (data not shown, see **section 3.4**). Ligands stabilize iGluR LBDs, which tend to aggregate in their apo forms, so all re-folding attempts were done both in the absence and presence of the IR84a ligand phenylacetaldehyde, which could help stabilize the LBD. However, phenylacetaldehyde is water insoluble and great difficulty was encountered working with phenylacetaldehyde in aqueous solutions. Due to this, further experiments with isolated IR84a in solution with phenylacetaldehyde were not conducted.

5.2.3 Expression and purification of SUMO-IR75a and SUMO-IR8a LBD in *E. coli.*

SUMO tags improve protein solubility [7]. Failure to express and re-fold IR84a prompted us to attempt expression of IR75a and IR8a LBDs with a SUMO tag. Expression of SUMO-IR75a/SUMO-IR8a was carried out in largely the same way as IR84a (see **section 3.4**). No detectable protein was present in the soluble fraction, but a considerable amount was present in the insoluble fraction (**Fig. 5-4**). Denatured SUMO-IR75a and SUMO-IR8a was purified on a nickel column in a denaturing buffer.

The denaturing buffer was slowly exchanged for a "stabilizing buffer" containing 10 mM propionic acid (IR75a only) and 1 M L-arginine and 15% glycerol, which stabilize proteins and prevent aggregation [9, 10]. Both proteins tolerated the stabilization buffer and was then dialyzed into a buffer containing only 5% glycerol and no L-arginine. At this step, ~65% of the protein aggregated, which marked an improvement over previous attempts to work in solution without GdnHCl, which always resulted in 100% aggregation. The remaining SUMO-IR8a/SUMO-IR75a was then incubated with the SUMO-specific protease (SENP), where detectable cutting took place (**Fig. 5-5** for IR75a and **Fig. 5-6** for IR8a). However, the protein aggregates (both uncut and cut with SENP) when dialyzed into solutions not containing any L-arginine or glycerol (data not shown). These results suggest IR8a and IR75a are stable in 5% glycerol solutions.

5.2.4 Co-expression of SUMO-IR75a and IR8a LBDs from a Duet plasmid.

It has been shown that co-receptors subunits stabilize odor-specific subunits in cellular environments [6]. The LBD layer of assembled IR tetramers are likely dimer-of-dimers in analogy to iGluRs, and it is possible that isolated LBDs will be stabilized by their dimer partners; IR LBDs may have aggregation-prone hydrophobic patches where their dimer partners are likely to bind. We co-expressed IR75a and IR8a from the

RSF-Duet1 plasmid. A Western blot indicates unsuccessful expression, with bands appearing only near 15 kDa, and not near the expected 30 kDa range **Fig. 5-7**. Repeated attempts to rectify this issue were not successful.

5.2.5 Expression of IR8a LBD in a baculovirus insect cell expression system.

Failure to express soluble IR84a, IR75a, and IR8a in *E. coli* led us to try IR expression in a baculovirus insect cell expression system. A recombinant bacmid was generated with the IR8a LBD, and trial expression was carried out in Hi5 and Sf9 cells. Expression was not observed in Sf9 cells (data not shown), but was observed in Hi5 cells (**Fig. 5-8**). The Western blot shown in **Fig. 5-8** indicates that the majority of IR8a partitioned to the insoluble fraction, but there is a faint band is visible in the soluble fraction.

5.3 Discussion

Revealing the atomic structures of IRs will help us to understand the molecular mechanism of olfaction in insects as well as the evolutionary history of ionotropic glutamate receptors. Structures could also aid in the rational design of insect repellents that target IR proteins. Here, I describe initial attempts to design and express isolated LBDs for IR8a, IR25a, IR75a, and IR84a in *E. coli* and a baculovirus insect cell expression system.

In *E. coli*, we attempted to express IR84a alone. The protein partitions to the insoluble fraction, and resists re-folding attempts. Additionally, its ligand, phenylacetaldehyde, is insoluble in water. Due to this, and the fact that iGluR LBDs are generally unstable in the absence of ligand, future work with isolated IR84a in aqueous solution is not recommended unless a more soluble ligand is used or IRs are generally found to be soluble without ligands.

IR8a and IR75a were expressed with N-terminal SUMO tags to enhance solubility. Similar to IR84a, these proteins partition to the insoluble fraction and resist refolding. There was some success preventing aggregation in the presence of L-arginine and glycerol, and even successful cutting with the SUMO-specific SENP protease, but the proteins resisted transfer out of solutions with aggregation inhibitors. Future work here might work to optimize buffer conditions that stabilize IRs in solutions with less than 5% glycerol.

IR8a and IR75a were also co-expressed from a Duet vector. Unfortunately, expression of full proteins was not observed by Western blot. It is possible that both proteins expressed, do not initially partition to the insoluble fraction, and are subsequently degraded by cellular machinery. The band observed near 15 kDa could correspond to the SUMO tag.

IR8a was expressed in baculovirus insect cell expression system. A small amount of protein was observed in the soluble fraction by Western blot. This is a promising avenue forward for future work. Also of note, all expression trials in insect cells discussed here were not done using the secretion machinery (i.e. trafficking through the ER and Golgi body). The IRs are natively shuffled through the secretion system, and sit exposed in an extracellular environment, so future work might consider using the honeybee melittin (HBM) secretion signal to create a more native-like IR protein folding environment.

Future studies might also re-visit construct design. For instance, odor-specific subunits have N-terminal regions of variable length that might improve solubility. In addition, because the S1 and S2 regions are discontinuous, the designed linkers designed here may not be sufficient to form properly folded IR LBDs. It may also be more feasible to skip isolated LBD expression entirely, and express IRs as full-length subunits.

5.4 Methods

5.4.1 Expression and purification of IR LBDs in *E. coli.*

IR84a LBD: The IR84a LBD was cloned into a pET28a vector from a pUC57 vector (GenScript), with the restriction enzymes Nde1 and EcoRI. Expression of IR84a was carried out in Rosetta(DE3) Competent Cells (MilliporeSigma). 1 L Terrific broth

(TB) was inoculated 20 mL IR84a starter culture and grown to $OD_{600} = 0.7$. Cultures were then cold shocked on ice for 20 minutes, when 400 μL IPTG was added to induce IR84a expression. Induction was carried out overnight at 4°C. Cells were harvested by spinning at 5000 RPM (H4000A 4 x 1 liter rotor, Sorvall) for 30 minutes at 4°C. The supernatant was discarded, and cells were suspended in a lysis buffer containing 150 mM NaCl, 20 mM Tris, pH 8. To this, 0.2 mM PMSF 10 μg/ml DNaseI were added. This mixture was stored on ice for 30 minutes, before cells were lysed in a microfluidizer. Lysed cells were then spun down at 15,000 RPM (SS-34 rotor, Sorvall) for 40 minutes at 4°C. Negligible protein was observed in the soluble fraction.

To clean up the insoluble fraction, it was suspended using a glass tissue homogenizer in a buffer containing 100 mM NaCl, 50 mM Tris, pH 8, 1% Triton, 10 mM DTT, 5 mM EDTA. This suspension was spun down at 15,000 RPM for 40 minutes at 4°C. The resulting insoluble fraction (insoluble fraction 2) was suspended in a buffer containing 100 mM NaCl, 50 mM Tris, pH 8, 10 mM DTT, 5 mM EDTA and again spun down at 15,000 RPM for 40 minutes at 4°C. The result is insoluble fraction 3, which was used for purifying and refolding IR LBDs on a nickel column. Purification of denatured IR LBDs from the nickel column is a standard nickel column purification, but in the presence of 6 M GdnHCl. Wash buffer: 6 M GdnHCl, 150 mM NaCl, 50 mM Tris, pH 8, 1 mM DTT, 10 mM imidazole. High Salt Wash buffer: Same as wash buffer, but with 500 mM NaCl. Elution buffer: Wash buffer with 400 mM imidazole.

SUMO-IR75a and SUMO-IR8a LBDs: IR75a and IR8a LBDs were cloned into pSUMO vectors with an N-terminal SUMO tag (provided by Cynthia Wolberger Laboratory) with the NsiI and XhoI restriction enzymes. Expression was carried out in an identical way to IR84a (see above).

SUMO-IR8a and IR75a from a RSF-Duet1 vector: SUMO-IR8a and IR75a were cloned into an RSF-Duet1 vector using BamHI-HF and EcoRI-HF for restriction enzymes for IR75a at multiple cloning site 1 (MCS1) and NdeI and XhoI for IR8a at MCS2. Successful insertion was verified using the primers ACYCDuetUP1 and DuetDOWN1 at MCS1 and DuetUP2 and T7-Term primers at MCS2. Cultures grown with kanamycin and chloramphenicol were induced with 400 μL IPTG at OD_{600} overnight at 4°C. The lysis procedure was the same as with IR84a (see above). The sample was run over a nickel column and eluted with 400 μM imidazole. A Western blot was run on samples.

5.4.2 Refolding IR LBDs from inclusion bodies

Flash dilution: Denatured IR solubilized in 6 M GdnHCl was added drop-wise to a rapidly spinning, room temperature 700 mL solution of 150 NaCl 50 mM Tris, pH 7, 1 mM DTT, and the 10 mM of LBD ligand (phenylacetaldehyde for IR84a or propionic acid for IR75a). Solution was then spun at 4000 RPM (Centrifuge 5810 R, Eppendorf) and filtered through 0.22 μM syringe-driven filter unit (Millipore) to remove aggregates.

Dropwise dilution: Denatured IR (0.5 mg/mL) solubilized in 6 M GdnHCl was initially dialyzed into a solution 6 M GdnHCl, 100 mM NaCl, 50 mM Tris, pH 8, 10 mM DTT, 5 mM EDTA, which was gradually replaced by a buffer of 150 mM NaCl, 10 mM Tris, pH 8, 1 mM DTT. Solutions were exchanged by a peristaltic pump (1 mL/min) for 48 hours. Solution was then spun at 4000 RPM (Centrifuge 5810 R, Eppendorf) and filtered through 0.22 μM syringe-driven filter unit (Millipore) to remove aggregates.

On column refolding: Denatured IR solubilized in 6 M GdnHCl was applied to a nickel column. The column was washed with a buffer not containing GdnHCl (150 mM NaCl, 10 mM Tris, pH 8, 1 mM DTT, 10 mM propionic acid (for IR75a only)) to promote refolding. An elution buffer with 400 mM imidazole was added to elute folded protein. This was followed by another wash in denaturing buffer, and then a denaturing buffer with 400 mM imidazole.

Dialysis into stabilizing buffers: The SUMO-IRs were suspended in a denaturing buffer (6 M GdnHCl, 100 mM NaCl, 50 mM Tris, pH 7, 10 mM DTT, 5 mM EDTA) which was substituted by dialysis into a stabilization buffer (1 M L-arginine, 20 mM HEPES, pH 7, 15% glycerol, 10 mM DTT). The proteins were then dialyzed into a solution without L-arginine, and lower concentrations of glycerol (150 mM NaCl, 20 mM HEPES, pH 7, 5% Glycerol, 1 mM DTT). 100 μg SUMO-specific protease (SENP) was added to separate the IRs from the SUMO tag and incubated for 1 day at 4°C and filtered through 0.22 μM syringe-driven filter unit (Millipore) to remove aggregates. Both cut and uncut samples were dialyzed into solutions not containing glycerol or L-arginine (150 mM NaCl, 20 mM HEPES, pH 7, 1 mM DTT, 10 mM propionic acid (IR75a only)).

5.4.3 Expression of the IR8a LBD in a baculovirus insect cell expression system.

Generating IR recombinant bacmids: IR LBD expression was carried out using the Bac-to-Bac Baculovirus Expression System (ThermoFisher). IR8a/IR25a/IR75a LBD genes were cloned into a pFastBacHT-A plasmid. IR-pFastBac plasmids were then transformed into DH10Bac *E. coli* which possess the bacmid DNA which will undergo recombination with IR-pFastBac to transfer IR LBD genes onto the bacmid. Transformed DH10Bac was plated on agar plates with 50 μg/mL kanamycin, 10 μg/mL tetracycline, 7 μg/mL gentamicin, 100 μg/mL Bluo-gal, and 40 μg/mL IPTG to select for positive recombination. White colonies indicate successful recombination. PCR using pUC-M13-F and pUC-M13-R primers was performed to verify positive recombination (see **Table 5.2**).

Generating high titer viral stock with recombinant bacmid DNA: 1-5 μg recombinant bacmid DNA (~500 ng/μL) was transfected into mid-log phase Sf9 insect cells in a 6-well plate, with ~1 million cells per plate. Transfected cells were incubated at 27°C for 96 hours without agitation to allow adhesion. P1 virus was collected from these samples and spun at 500 *g* for 5 min for minutes to remove cellular debris. 500 μL P1 virus was then used to infect 12 mL of mid-log phase Sf9 cells (1-2x10^6 cells/mL) in a rectangular cell culture flask and incubated without agitation at 27°C for 72 hours. P2 virus was collected and spun at 500 g for 5 minutes to remove debris. P3 stock was generated by infecting 100 mL of Sf9 cells with P2 virus, which again was incubated at 27°C for 72 hours and spun at 500 g for 5 minutes to remove debris.

Expression of IR8a into Sf9 and Hi5 insect cells: To express IR8a in Sf9 and Hi5 insect cells, 20mL P3 stock was added per 1 L of Sf9 and Hi5 cell cultures in mid-log phase (1-2x10^6 cells/mL). Hi5 cells were incubated at 27°C for 48 hours, and Sf9 cells were incubated 27°C for 72 hours. Cells were spun down at 5000 RPM for 30 minutes and resuspended in lysis buffer containing 300 mM NaCl, 50 mM Tris, pH 8, 5% glycerol, and 0.1% Triton X-100. To this, 0.2 mM PMSF, a protease inhibitor cocktail tablet (Roche), and 10 μg/ml DNaseI were added. Cells were incubated on ice for 30 minutes, and then lysed in a microfluidizer, and spun down at 15,000 RPM at 4°C for 45 minutes. Sample were taken from the supernatant and insoluble fraction were for Western blot.

5.4.4 Western blot of his-tagged IRs:

Western blotting was performed on his-tagged IR with a rabbit anti-His antibody (Cell Signaling Technology). Samples run on SDS-PAGE were transferred to a PVDF membrane (Bio-Rad) in a transfer buffer (25 mM Tris, pH 8.3, 192 mM glycine, 20% (v/v) methanol, 0.1% Triton X-100) at constant current (350 mA) for 1 hour. Blocking was carried out in a solution of 5% nonfat dry milk and PBS with 0.1% Triton X-100 at room temperature for 1 hour. The membrane was incubated with the primary rabbit anti-his antibody at 4°C overnight, with constant agitation. Following a wash, the membrane was then incubated with the HRP-linked anti-Rabbit IgG secondary antibody (Cell Signaling Technology) for 1 hour at room temperature. HRP activity was induced by incubating the membrane with SignalFire ECL Reagent (Cell Signaling Technology) for 1 minute at room temperature.

```
IR25a     DEEQMKNLTADTVYRIFTVVQAPFIM-----------------------------RDETA   458
GluN2A    YKSFSDCEPDDNHLSIVTLEEAPFVIVEDIDPLTETCVRNTVPCRKFVKINNSTNEGMNV   452


IR25a     PKGYKGYCIDLINEIAAIVHFDYTIQEVEDGKFGNMDENGQWNGIVKKLMDKQADIGLGS   518
GluN2A    KKCCKGFCIDILKKLSRTVKFTYDLYLVTNGKHGKK-VNNVWNGMIGEVVYQRAVMAVGS   511


IR25a     MSVMAEREIVIDFTVPYYDLVGITIMMQRPSSPSSLFKFLTVLETNVWLCILAAYFF-TS   577
GluN2A    LTINEERSEVVDFSVPFVE-TGISVMVSRSNGTVSPSAFLEPFSASVWVMMFVMLLIVSA   570


IR25a     FLMWIFDRWSPYSYQNNREKYKDDEEKREFNLKECLWFCMTSLT-PQGGGEAPKNLSGRL   636
GluN2A    IAVFVFEYFSPVGYNRNLAKGKA-PHGPSFTIGKAIWLLWGLVFNNSVPVQNPKGTTSKI   629


IR25a     VAATWWLFGFIIIASYTANLAAFLTVSRLDTPVESLDDLAKQYKILYAPLNGSSAMTYFE   696
GluN2A    MVSVWAFFAVIFLASYTANLAAFMIQEEFVDQVTGLSDKKFQRPHDYSPPFRFGTVPNGS   689


IR25a     RMSNIEQMFYEIWKDLSLNDSLTAVERSKLAVWDYPVSDKYTKMWQAMQEAKLPATLDEA   756
GluN2A    TERN--------------------------IRNNYPYMHQYMTRFNQRG-------VEDA   716


IR25a     VARVRNSTAATGFAFLGDATDIRYLQLTNCDLQVVG--EEFSRKPYAIAVQQGSHLKDQF   814
GluN2A    LVSLKT-GKLDAFIYDAAVLNYKAGRDEGCKLVTIGSGYIFASTGYGIALQKGSPWKRQI   775


IR25a     NNAILTLLNKRQLEKLKEKWWKNDEALAKCDKPEDQSDGISIQNIGGVFIVIFVGIGMAC   874
GluN2A    DLALLQFVGDGEMEELETLWLTGI---CHNEKNEVMSSQLDIDNMAGVFYMLAAAMALSL   832
```

Fig. 5-1: **IR LBD construct design.** Alignment between IR25a and GluN2A surrounding the LBD segments. In yellow: sequence of the GluN2A LBD as in PDB: 2A5S. In red: sequence included in construct built for IR25a. Residues underlined and italicized are either known to be or predicted to be inserted in the plasma membrane.

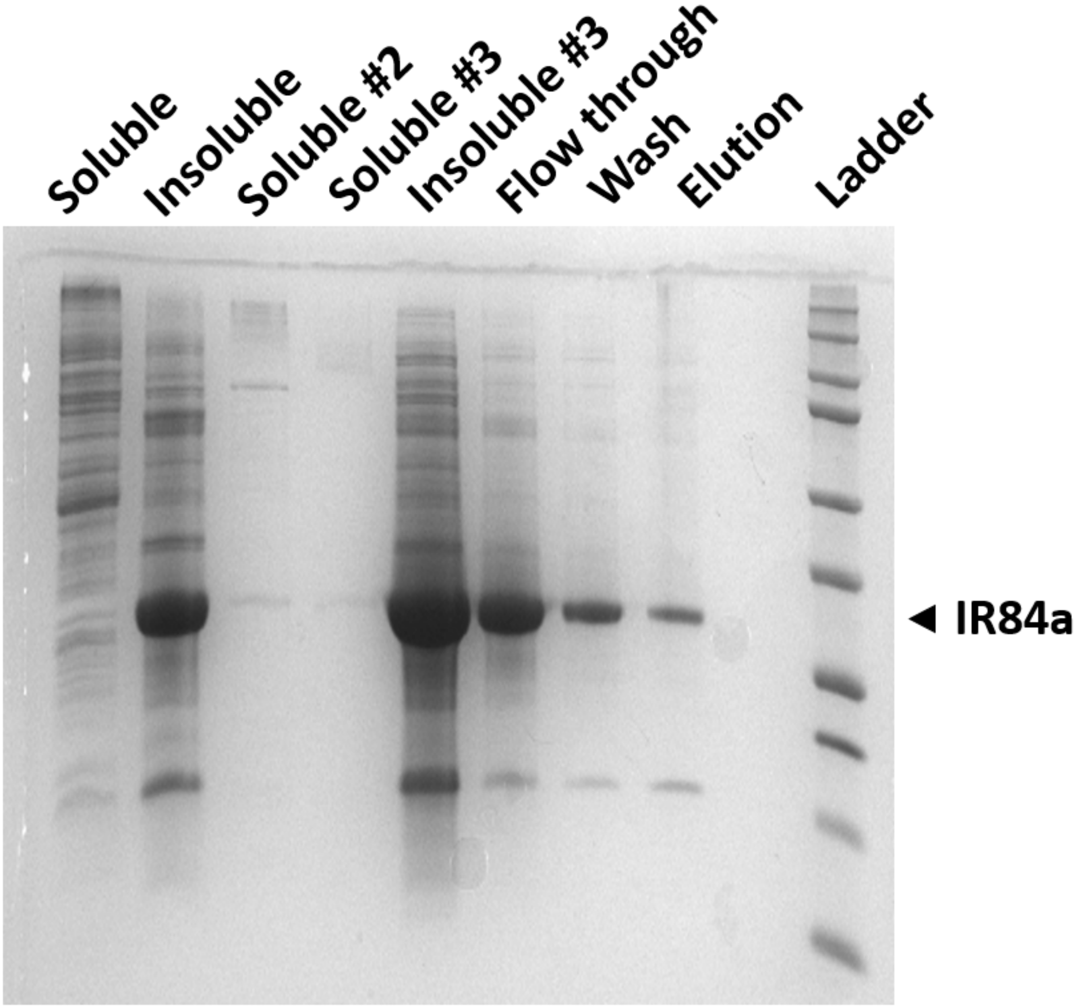


Fig. 5-2: **Purification of denatured IR84a on a nickel column.** SDS-PAGE results of IR84a expressed in *E coli*. Soluble and insoluble fractions (*lanes 1-2*). *Lanes 3-4* show clean up of insoluble fraction through repeated suspension and spin downs. *Lanes 6-8* show purification of denatured IR on a nickel column.

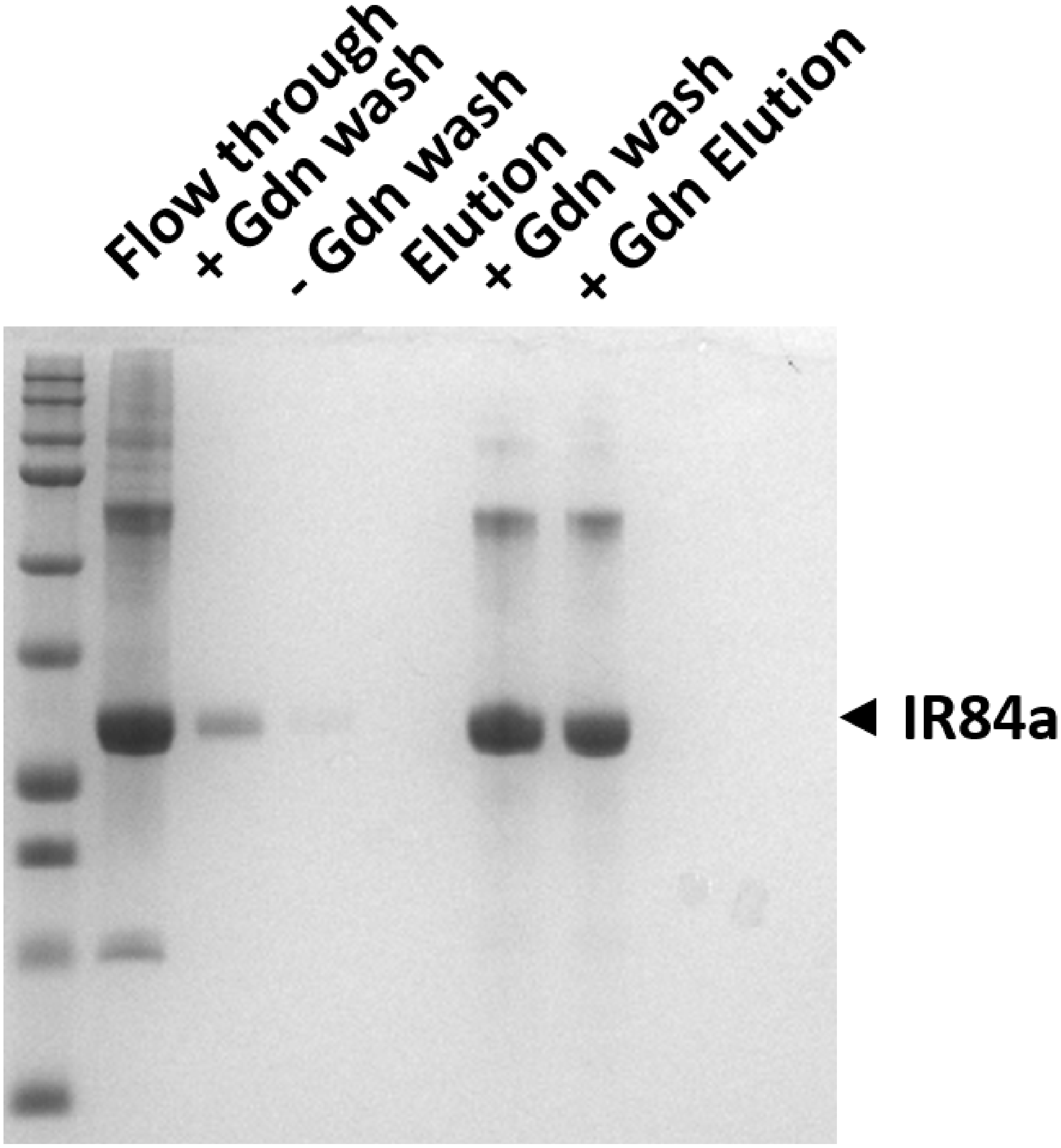


Fig. 5-3: **Refolding IR84a on a nickel column.** Denatured IR84a was applied to a nickel column, *lane 2* shows the flow through. *Lanes 4-5* show a wash without GdnHCl, to prompt protein folding, followed by an attempted elution of folded IR84a. Lanes 6-7 show an additional GdnHCl wash, followed by an elution. This shows that IR84a likely aggregates on the column, as it is not eluted in the absence of denaturant.

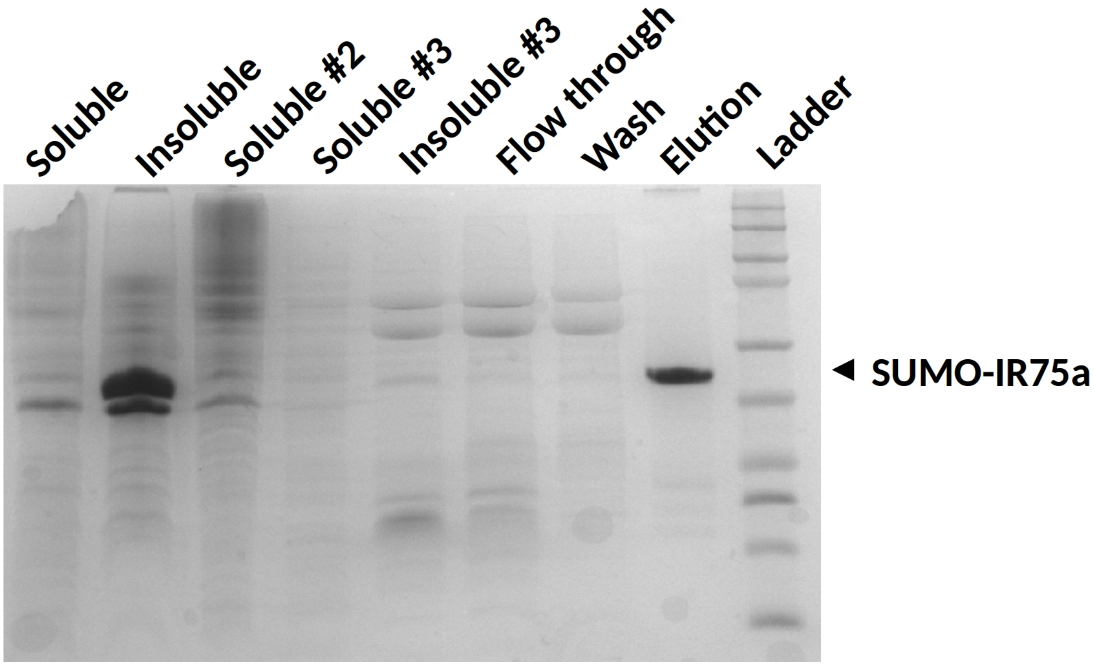


Fig. 5-4: **Purification of denatured SUMO-IR75a on a nickel column.** Soluble and insoluble fractions in *lanes 1-2.* Clean up of the insoluble fraction *lanes 3-5.* Purification on a nickel column in *lanes 6-8.*

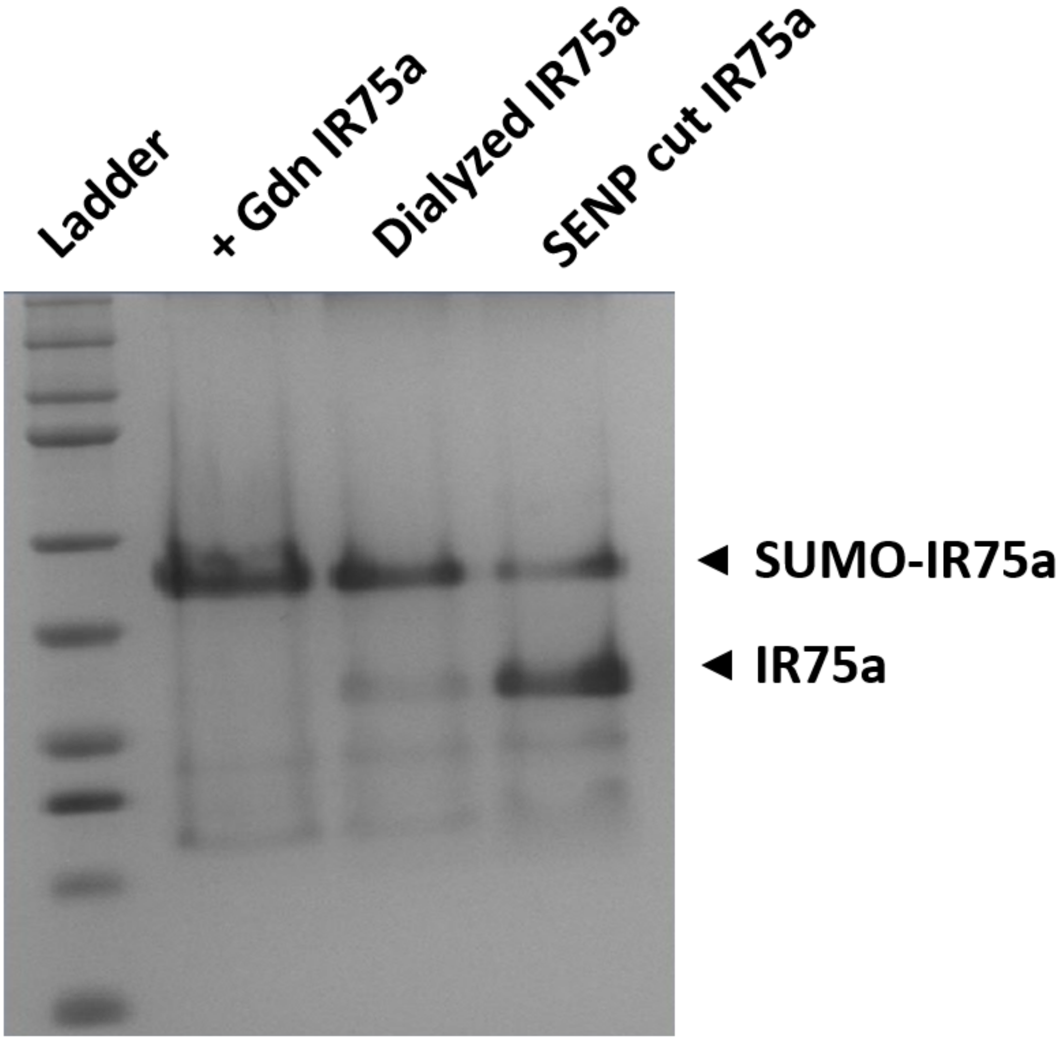


Fig. 5-5: **Dialysis of SUMO-IR75a into stabilization buffer and SENP protease cleavage.** *Lane 2* shows SUMO-IR75a in a 6 M GdnHCl denaturing buffer. *Lane 3* shows SUMO-IR75a after is has been transferred to a stabilization buffer containing 5% glycerol. *Lane 4* shows successful SUMO-IR75a cleavage by SENP protease indicated IR75a LBD does not aggregate on its own in solutions containing glycerol.

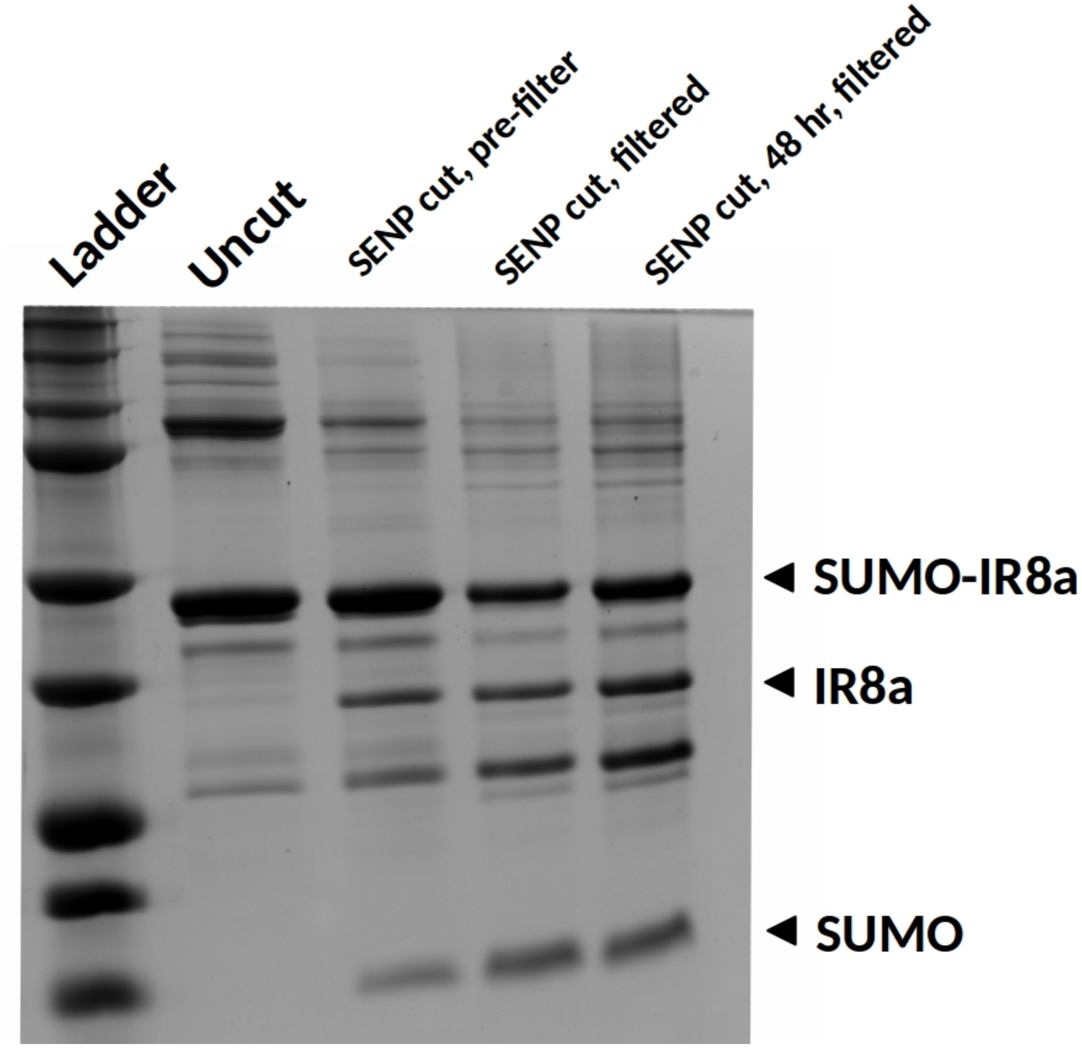


Fig. 5-6: **SENP protease cleavage of SUMO-IR8a.** *Lane 2* shows uncut SUMO-IR8. *Lane 4* shows cut and filtered SUMO-IR8. The bands labeled IR8a and SUMO become visible only after cutting with SENP. *Lane 5* shows cut SUMO-IR8a after 48 hours, indicated resistance to aggregation over time in 5% glycerol solution.

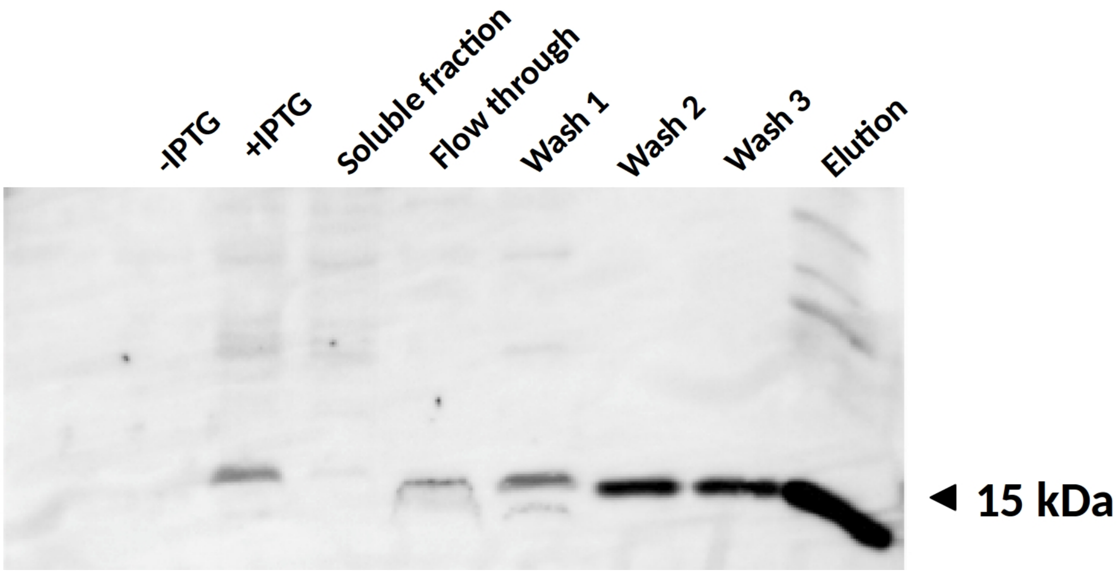


Fig. 5-7: **Western blot from co-expression of SUMO-IR8a and IR75a from RSF-Duet1 vector.** Western blot from nickel column purification of SUMO-IR8a and IR75a co-expression. A consistent band near 15 kDa is observed, which is possibly the SUMO protein. The presence of IR LBDs is not detected (expected band near 45-50 kDa for SUMO-IR8a).

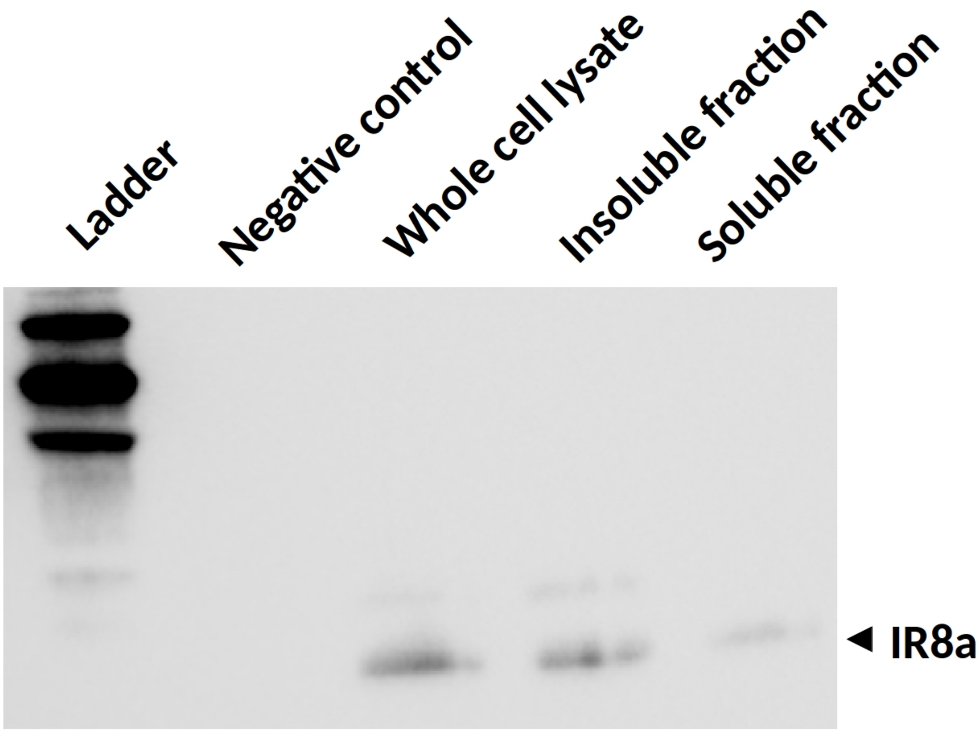


Fig. 5-8: **Expression of IR8a LBD in a baculovirus insect cell expression system.** Anti-his Western blot of IR8a expression. Expression is evident in whole cell lysate (*lane 3*). Most protein partitions to the insoluble fraction, however there is a faint band in the soluble fraction *lanes 4-5*.

IR8a:ARRFFRIGTAESIPWSYLRREEGTGELIRDRSGLPIWEGYCIDFIIRL
SQKLNFEFEIVAPEVGHMGELNELGEWDGVVGDLVRGETDFAIAALK
MYSEREEVIDFLPPYYEQTGISIAIRKGTPVQSLEQLARQSRINYTV
VKDSDTHQYFVNMKFAEDTLYRMWKELALNASKDFKKFRIWDYPIKE
QYGHILLAINSSQPVADAKEGFANVDAHENADYAFIHDSAEIKYEIT
RNCNLTEVGEVFAEQPYAVAVQQGSHLGDELSYAILELQKDRFFEEL
KAKYWNQSNLPN

IR25a:ADTVYRIFTVVQAPFIMRDETAPKGYKGYCIDLINEIAAIVHFDYTIQ
EVEDGKFGNMDENGQWNGIVKKLMDKQADIGLGSMSVMAEREIVIDF
TVPYYDLVGITIMMQRPSSGTVESLDDLAKQYKILYAPLNGSSAMTY
FERMSNIEQMFYEIWKDLSLNDSLTAVERSKLAVWDYPVSDKYTKMW
QAMQEAKLPATLDEAVARVRNSTAATGFAFLGDATDIRYLQLTNCDL
QVVGEEFSRKPYAIAVQQGSHLKDQFNNAILTLLNKRQLEKLKEKWW
KNDEALAKCDK

IR75a:TDVVLRVATVVTQRPLTLSDDELIRFLSQENDTHIDSLARFGFHLTLI
LRDLLHCKMKFIFSDSWSKSDVVGGSVGAVVDQTADLTATPSLATEG
RLKYLSAIIETGFFRSVCIFRTGTKIKTMRQLAESSLTVGLEPLPFT
KSYLNYSRLPEIHLFIKRKIESQTQNPELWLPAEQGVLRVRDNPGYV
YVFETSSGYAYVERYFTAQEICDLNEVLFRPEQLFYTHLHRNSTYKE
LFRLRFLRILETGVYRKQRSYWVHMKLHCVA

IR84a:FLNVTLIGSTVLTEKPPGFGDMEYLADDKQLQQLDPMQRKTYQLFQLV
ERMFNLSLAISLTDKWGELLDNGSWSGVMGQVTSREADFAVCPIRFV
LDRQPYVQYSAVLHTQNIHFLFRHGTIVNLQALYDSNLAIGMENISY
NFPIFTNTSNQLVRDVYVKKICKSGEHNIMSLQQGAERIIQGRFAFH
TAIDRMYRLLLELQMDEAEFCDLQEVMFNLPYDSGSVMPKGSPWREH
LAHALLHFRATGLLQYNDKKWMVRRPDCSL

Table 5.1: **Sequences of IR8a, IR25a, IR75a, and IR84a LBDs.**

Primers for cloning IR LBDs into pFastBac-HTA plasmid:

IR8a-F: TAAGCGAATTCGCGCGTCGTTTCTTTCGTATCGGCAC
IR8a-R: AAGCCTCGAGTTAGTTCGGCAGGTTGCTCTGGTTCCAATAC
IR75a-F: TAAGCGAATTCACCGATGTGGTGCTGCGTGTTGCG
IR75a-R: AAGCCTCGAGTTACGCCACGCAATGCAGTTTCATATGC
IR8a-Frmshft-F: CAGGGCGCCATGGATGCGCGTCGTTTCTTTC
IR8a-Frmshft-R: GAAAGAAACGACGCGCATCCATGGCGCCCTG
IR25a-Frmshft-F: TTCAGGGCGCCATGGATGCCGATACGGTCTACAG
IR25a-Frmshft-R: CTGTAGACCGTATCGGCATCCATGGCGCCCTGAA
IR75a-Frmshft-F: CAGGGCGCCATGGATACCGATGTGGTGCTGCG
IR75a-Frmshft-R: CGCAGCACCACATCGGTATCCATGGCGCCCTG

Primers to verify pFastBac-HTA inserts:

pUC-M13-F: CCCAGTCACGACGTTGTAAAACG
pUC-M13-R: AGCGGATAACAATTTCACACAGG

Primers for IR transfer from pUC57 to RSF-Duet1 vector:

Duet-IR75a-F: CATGGATCCACCGATGTGGTGCTGCGTGTTGCG
Duet-IR75a-R: TAAGCGAATTCTTACGCCACGCAATGCAGTTTCATATGC
Duet-IR8a-F: TACCATATGGCGCGTCGTTTCTTTCGTATCGGC
Duet-IR8a-R: AAGCCTCGAGTTAGTTCGGCAGGTTGCTCTGGTTCCAATAC
Duet-IR75a-frmshft-F: CATCACCACAGCCAGGATACCGATGTGGTGCTGCGT
Duet-IR75a-frmshft-R: ACGCAGCACCACATCGGTATCCTGGCTGTGGTGATG

Primers for verifying RSF-Duet1 vector inserts:

ACYCDuetUP1: GGATCTCGACGCTCTCCCT
DuetDOWN1: GATTATGCGGCCGTGTACAA
DuetUP2: TTGTACACGGCCGCATAATC

Table 5.2: Primers used to clone IR LBDs.

References

[1] Silbering, A. F. and Benton, R. Ionotropic and metabotropic mechanisms in chemoreception:'chance or design'? *EMBO reports*, 11(3):173–179, 2010.

[2] Benton, R., Vannice, K. S., Gomez-Diaz, C., and Vosshall, L. B. Variant ionotropic glutamate receptors as chemosensory receptors in Drosophila. *Cell*, 136(1):149–162, 2009.

[3] Ni, L., Klein, M., Svec, K. V., Budelli, G., Chang, E. C., Ferrer, A. J., Benton, R., Samuel, A. D., and Garrity, P. A. The ionotropic receptors IR21a and IR25a mediate cool sensing in Drosophila. *Elife*, 5:e13254, 2016.

[4] Enjin, A., Zaharieva, E. E., Frank, D. D., Mansourian, S., Suh, G. S., Gallio, M., and Stensmyr, M. C. Humidity sensing in Drosophila. *Current Biology*, 26(10):1352–1358, 2016.

[5] Abuin, L., Bargeton, B., Ulbrich, M. H., Isacoff, E. Y., Kellenberger, S., and Benton, R. Functional architecture of olfactory ionotropic glutamate receptors. *Neuron*, 69(1):44–60, 2011.

[6] Ai, M., Blais, S., Park, J.-Y., Min, S., Neubert, T. A., and Suh, G. S. Ionotropic glutamate receptors IR64a and IR8a form a functional odorant receptor complex in vivo in Drosophila. *Journal of Neuroscience*, 33(26):10741–10749, 2013.

[7] Kuo, D., Nie, M., and Courey, A. J. SUMO as a solubility tag and in vivo cleavage of SUMO fusion proteins with Ulp1. In *Protein Affinity Tags*, pages 71–80. Springer, 2014.

[8] Singh, A., Upadhyay, V., Upadhyay, A. K., Singh, S. M., and Panda, A. K. Protein recovery from inclusion bodies of Escherichia coli using mild solubilization process. *Microbial cell factories*, 14(1):41, 2015.

[9] Gekko, K. and Timasheff, S. N. Mechanism of protein stabilization by glycerol: preferential hydration in glycerol-water mixtures. *Biochemistry*, 20(16):4667–4676, 1981.

[10] Tsumoto, K., Umetsu, M., Kumagai, I., Ejima, D., Philo, J. S., and Arakawa, T. Role of arginine in protein refolding, solubilization, and purification. *Biotechnology progress*, 20(5):1301–1308, 2004.

Chapter 6 - Concluding Remarks

This dissertation examines the conformational dynamics of various clamshell-like ligand binding domains (LBDs) primarily using computational and experimental approaches. The clamshell-like LBDs all belong to a large family of proteins that descend from the periplasmic binding proteins in gram negative bacteria. Generally, these proteins bind ligands which stabilize a closed cleft conformation over an open cleft conformation. Structural characterization of these LBDs began over 40 years ago with the crystal structure of L-arabinose at 5 Å. Since then, this field has remained remarkably active, and has potential applications to drug discovery, biosensors, optogenetics, enzyme and protein design, and more.

Chapter 1 introduces PBPs and the mammalian LBDs for ionotropic glutamate receptors (iGluRs), metabotropic glutamate receptors (mGluRs), and insect ionotropic receptors (IRs). I review the function, structure, and dynamics of the LBDs as well as the full receptors in the case of ionotropic glutamate receptors (iGluRs), metabotropic glutamate receptors (mGluRs), and the ionotropic receptors (IRs). The iGluRs and mGluRs are required in mammalian central nervous system, and play an essential role in regulating the transmission of electrical signal between neurons. Uncovering the structural and dynamical features of these proteins will help to build a more accurate model of the mammalian synapse. The IRs are olfactory receptors in insects that are homologous to iGluRs. A structural model of an IR will help to reveal the mechanism of odor detection in insects, as well as to aid our understanding of divergent evolutionary pathway of iGluRs and IRs.

Chapter 2 introduces the primary computational method utilized in later chapters. Namely, there is a brief introduction to molecular dynamics simulation and umbrella sampling, the tool we use to compute PMFs, or free energy landscapes of LBDs. This chapter also gives a practical overview of how to set up an LBD system for umbrella sampling using the order parameters we define for most of our systems.

Chapter 3 presents the first PMFs for a kainate receptor LBD. We find that the apo GluK2 LBD exhibits high conformational flexibility that spans open and closed conformations. We also find that the glutamate-bound GluK2 LBD is more flexible than the LBDs of other iGluR subtypes. Mutations that disrupt interactions across the binding cleft have a variable effect on closed state stability: one mutation appears to have no effect at all, while the other has a significant effect. Finally, we show that theoretical SAXS from our computational ensemble of glutamate-bound GluK2 better fit with experimental results obtained in the presence of saturating glutamate. Significant effort was spent trying to perform SAXS on apo GluK2, as well as to grow crystals for X-ray crystallography. Apo GluK2 is prone to aggregation and is sensitive to X-ray beams. Characterizing apo GluK2 is left to a future member of the lab. One general question remains: what give rise to the unique properties of iGluR LBDs? Here we present some analysis looking at bulk and surface electrostatics, but a more thorough investigation is needed. It is also unclear how these properties translate to the full length receptor. Future computational studies of kainate receptors might involve more complicated setups: LBD dimers, ATD-LBD systems, full-length tetramers.

Chapter 4 presents PMFs for mGluR3, which are the first PMF presented for any clamshell-like domain descended from the Class I PBPs. We show that mGluR3 monomers are highly flexible, but that flexibility is lost in a dimer system where both LBDs are glutamate-bound and at least one LBD is constrained closed. This suggests an allosteric effect of dimerization, and we observed a change in the dynamics of a single helix in the lower lobe that might account for the change observed in the PMF. We also present the PMF along an order parameter corresponding to the transition between inactive and active dimer states. The PMF is noisy and can likely be improved with either more sampling, tuning the force constant, or a change of order parameters. Future studies on mGluRs might look at other subtypes, ligand binding pathways, and systems of the full-length receptor.

Chapter 5 is a first attempt at expressing and purifying IR LBDs. Generally, the IR LBDs I designed appear to be insoluble. We attempted a wide array of strategies: different subtypes, refolding from inclusion body, adding a solubilizing tag, co-expression, and baculovirus insect cell expression. The baculovirus insect cell expression has some promising results, with a small amount of protein appearing in the soluble fraction. This result needs to be followed up on in future work. However, it might be worth revisiting LBD construct design to see if there is anything that can be done to improve upon current designs. If soluble LBD is obtained, ligand-binding assays can be done in addition to structural work. Finally, it might be easier to skip solving the LBD structure entirely, and work only in the context of a full-length receptor.

Curriculum vitae

TYLER WIED

PERSONAL INFORMATION

address 520 Park Avenue, Apt 610, Baltimore, MD 21201

email twied1@jhmi.edu

phone (920) 634 9618

EDUCATION

2012-Present Johns Hopkins University School of Medicine

PhD in Biophysics

National Science Foundation Graduate Research Fellow
Thesis: *Structural characterization and conformational dynamics of clamshell-like ligand-binding domains*

2008-2012 University of Wisconsin-Madison

Bachelors in Biochemistry

Graduated with Distinction and Honors in Research
Thesis: *Thermodynamics of solute effects on triple stranded DNA formation*

RESEARCH EXPERIENCE

2012–Present Graduate Research Assistant – School of Medicine

Johns Hopkins University

Investigated how receptors at neuronal synapses undergo conformational change via a combination of computational and experimental approaches. Published research in peer-reviewed journals and presented work at major scientific conferences. Managed undergraduate and graduate students within the lab. Awarded NSF Graduate Research Fellowship, total award amount $115,000.
Advisor: Albert Y. Lau, PhD · (443) 287 4528 · alau@jhmi.edu

2010–2012 Undergraduate Research Assistant – Department of Biochemistry

University of Wisconsin-Madison

Completed a senior thesis on the thermodynamics of triple-stranded DNA formation. The goal of the broader research program is to build a quantitative framework to predict the effect of solutes on macromolecular processes. Responsible for optimizing experimental conditions as well as data collection and analysis.
Advisor: M. Thomas Record, Jr., PhD · (608) 262 5332 · mtrecord@wisc.edu

2009–2011 Undergraduate Research Assistant – Department of Zoology

University of Wisconsin-Madison

Published research on the effects of lithium on naturally hyperactive mice in the development of a mouse model for mania. Responsible for carrying out behavioral tests, data analysis, and basic mouse care.
Advisor: Stephen Gammie, Jr., PhD · (608) 262 3457 · scgammie@wisc.edu

PUBLICATIONS

Wied T, and Lau AY. Conformational dynamics of a metabotropic glutamate receptor ligand-binding domain. *In preparation*.

Wied T, Chin AC, Lau AY. High conformational variability in the GluK2 kainate receptor ligand-binding domain. *Accepted at Structure*.

Yu A, **Wied T**, Belcher J, and Lau AY. Computing conformational free energies of iGluR ligand binding domains. In *Ionotropic Glutamate Receptor Technologies*, pp. 119-132. Humana Press, New York, NY, 2016.

Scotti MA, Lee G, Stevenson SA, Ostromecki AM, **Wied TJ**, Kula DJ, Gessay GM, Gammie SC. (2011). Behavioral and pharmacological assessment of a potential new mouse model for mania. *Physiol Behav*. 103(3-4), 376-383. PMID: 21397618.

TEACHING AND MENTORING EXPERIENCE EXPERIENCE

2015–Present · Mentor to several undergraduate and junior graduate students in the Lau lab.

2013–2015 · Teaching Assistant, Biochemical and Biophysical Principles, Johns Hopkins University School of Medicine.

2014 · Teaching Assistant, Macromolecular Structure and Analysis, Johns Hopkins University School of Medicine.

2011–2012 · Fun with Chemistry Camps Group Leader, Institute for Chemical Education, University of Wisconsin-Madison.

2010–2012 · Tutor, Greater University Tutoring System, University of Wisconsin-Madison.

OTHER INFORMATION

Awards

2013-2017 · National Science Foundation Graduate Research Fellowship

2011 · Albert J and Adelaide Riker Scholarship

2010 · Walter H. Ebeling Scholarship

Presentations & Talks

2018 · Conformational dynamics of a class C G protein coupled receptor. *Talk at Dept. of Biophysics and Biophysical Chemistry Intradepartmental Seminar Series*. Johns Hopkins University, Baltimore, MD.

2017 · Updates on glutamate receptor simulations: from comb jellies to umami. *Talk at Dept. of Biophysics and Biophysical Chemistry Intradepartmental Seminar Series*. Johns Hopkins University, Baltimore, MD.

2017 · Free energy landscapes of metabotropic glutamate receptor ligand-binding domains. *Poster at The 61st Annual Meeting of the Biophysical Society*. New Orleans, LA.

2016 · Conformational dynamics of the GluK2 ligand-binding domain. *Talk at the BCMB Student Colloquium.* Johns Hopkins University, Baltimore, MD.

2016 · Conformational dynamics of the GluK2 ligand-binding domain. *Poster at the 60th Annual Meeting of the Biophysical Society*. Los Angeles, CA.

2015 · Conformational dynamics of the GluK2 ligand-binding domain. *Poster at The 59th Annual Meeting of the Biophysical Society*. Baltimore, MD.

September 17, 2018